Preprint ANL/MCS-P864-1200

ON USING THE ELASTIC MODE IN NONLINEAR PROGRAMMING APPROACHES TO MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS

MIHAI ANITESCU*

Abstract. We investigate the possibility of solving mathematical programs with complementarity constraints (MPCCs) using algorithms and procedures of smooth nonlinear programming. Although MPCCs do not satisfy a constraint qualification, we establish sufficient conditions for their Lagrange multiplier set to be nonempty. MPCCs that have nonempty Lagrange multiplier sets and that satisfy the quadratic growth condition can be approached by the elastic mode with a bounded penalty parameter. In this context, the elastic mode transforms MPCC into a nonlinear program with additional variables that has an isolated stationary point and local minimum at the solution of the original problem, which in turn makes it approachable by sequential quadratic programming algorithms. One such algorithm is shown to achieve local linear convergence once the problem is relaxed. Under stronger conditions, we also prove superlinear convergence to the solution of an MPCC using an adaptive elastic mode approach for a sequential quadratic programming algorithm recently analyzed in an MPCC context by Fletcher and al. [18]. Our assumptions are more general since we do not use a critical assumption from that reference. In addition, we show that the elastic parameter update rule will not interfere locally with the super linear convergence once the penalty parameter is appropriately chosen.

Key words. Nonlinear Programming, Elastic Mode, SQP, MPEC, MPCC, complementarity constraints.

AMS subject classifications. 90C30, 90C33, 90C55,49M37, 65K10

1. Introduction.

Complementarity constraints can be used to model numerous economics or engineering applications [30, 36]. Solving optimization problems with complementarity constraints may prove difficult for classical nonlinear optimization, however, given that, at a solution x^*, such problems cannot satisfy a constraint qualification [30, Chapter 3]. As a result, algorithms based on the linearization of the feasible set, such as sequential quadratic programming (SQP) algorithms, may fail because feasibility of the linearization can no longer be guaranteed in a neighborhood of the solution [30].

Several methods have been recently proposed to accommodate such problems. Nonsmooth and disjunctive programming approaches [31, 30, 36] can be used to successfully solve MPCC. However, for certain problems they may take a number of steps that is exponential in the size of the problem. For special cases, such that linear data functions, or bi-level optimization, other succesfull approaches have been defined [22, 30].

In this work we investigate an elastic mode SQP approach for MPCC. The elastic mode is a standard technique of approaching infeasible subproblems by relaxing the constraints and introducing a differentiable penalty term in the objective function [24].

Here we use the framework from [40] to determine sufficient conditions for MPCCs

* Mathematics and Computer Science Division, Argonne National Laboratory, 9700 South Cass Ave., Argonne, Illinois, 60439, (anitescu@mcs.anl.gov).

to have nonempty Lagrange multiplier sets. As in [40], the first- and second-order optimality properties of an MPCC are compared with the similar properties of two nonlinear programs that involve no complementarity constraints.

Note that the elastic mode formulation is different from regularization approaches [25, 30, 41, 28] (though the latter reference introduces an elastic mode relaxation, but not with respect to all nonlinear constraints, and, in particular, not with respect to the complementarity constraints). The regularization approaches create a family of nonlinear programs that are regular (in the sense that they satisfy a constraint qualification at the solution) for the interior of the domain of the regularization parameter. However, the solution of the regularized roblems is different from the one of the original MPCC for any value of the parameter in the interior of the domain. The solution of the MPCC is obtained in the limit of the regularization parameter going towards the boundary of its domain. By contrast, under certain conditions to be developed in this work, the elastic mode formulation transforms the original MPCC into a regular nonlinear program with the same solution as the MPCC.

We also note that SQP approaches have been applied before to MPCC, in connection to regularization methods [21]. However, this is the first work that analyzes in detail issues connected to applying an elastic mode SQP approach to solving MPCC.

The paper is structured as follows.

- In Section 1 we review the relevant nonlinear programming concepts.
- In Section 2 we discuss sufficient conditions for MPCC to have a nonempty Lagrange multiplier set, in spite of not satisfying a constraint qualification at any point.
- In Section 3 we show that the elastic mode applied to an instance of the MPCC class will retrieve a local solution of the problem for a finite value of the penalty parameter, a point which is supported by several numerical examples.
- In Section 4 we prove that an adaptive elastic mode approach built around an algorithm recently analyzed in Fletcher and al. [18] in the MPCC context will result in super linear convergence near the solution of an MPCC under assumptions weaker than in [18]. Specifically, here we do not assume that the iterates are either feasible or satisfy the complementarity constraints for the unrelaxed problem. If the sequence produced by the algorithm is assumed to converge, we show that superlinear convergence follows from even weaker assumptions about the signs of the relevant multipliers. In addition, we show that the elastic parameter update rule will not affect locally the super linear convergence once the penalty parameter is appropriately chosen.

1.1. Optimality Conditions for General Nonlinear Programming. We review the optimality conditions for a general nonlinear program

$$\min_x \tilde{f}(x) \quad \text{subject to } \tilde{g}(x) \leq 0,\ \tilde{h}(x) = 0. \tag{1.1}$$

Here $\tilde{g} : \mathcal{R}^n \to \mathcal{R}^m$, $\tilde{h} : \mathcal{R}^n \to \mathcal{R}^r$. We assume that $\tilde{f}$, $\tilde{g}$, and $\tilde{h}$ are twice continuously differentiable.

We call x a stationary point of (1.1) if the Fritz-John condition holds: There exist multipliers $0 \neq \tilde{\lambda} = (\tilde{\lambda}_0, \tilde{\lambda}_1, \ldots, \tilde{\lambda}_{m+r}) \in \mathcal{R}^{m+r+1}$, such that

$$\nabla_x \mathcal{L}(x, \tilde{\lambda}) = 0,\ \ \tilde{h}(x) = 0;\ \ \tilde{\lambda}_i \geq 0,\ \ \tilde{g}_i(x) \leq 0, \quad \text{for } i = 1, 2, \ldots, m;\ \ \sum_{i=1}^m \tilde{\lambda}_i \tilde{g}_i(x) = 0. \tag{1.2}$$

Here $\mathcal{L}$ is the Lagrangian function

$$\mathcal{L}(x, \tilde{\lambda}) = \tilde{\lambda}_0 \tilde{f}(x) + \sum_{i=1}^{m} \tilde{\lambda}_i \tilde{g}_i(x) + \sum_{j=1}^{r} \tilde{\lambda}_{m+j} \tilde{h}_j(x). \tag{1.3}$$

A local solution x^* of (1.1) is a stationary point [37]. We introduce the sets of generalized Lagrange multipliers

$$\Lambda^g(x) = \left\{ 0 \neq \tilde{\lambda} \in \mathcal{R}^{m+r+1} \mid \ \tilde{\lambda} \text{ satisfies (1.2) at } x \right\}, \tag{1.4}$$

$$\Lambda_1^g(x) = \left\{ \tilde{\lambda} \in \Lambda^g(x) \mid \tilde{\lambda}_0 = 1 \right\}. \tag{1.5}$$

The set of active inequality constraints at a stationary point x is

$$\tilde{\mathcal{A}}(x) = \left\{ i \in \{1, 2, \ldots, m\} \mid \tilde{g}_i(x) = 0 \right\}. \tag{1.6}$$

The set of inactive inequality constraints at x is the complement of $\tilde{\mathcal{A}}(x)$:

$$\tilde{\mathcal{A}}^c(x) = \{1, 2, \ldots, m\} - \tilde{\mathcal{A}}(x). \tag{1.7}$$

With this notation, the complementarity condition from (1.2), $\sum_{i=1}^{m} \tilde{\lambda}_i g_i(x) = 0$, becomes $\tilde{\lambda}_{\tilde{\mathcal{A}}^c(x)} = 0$.

If certain regularity conditions hold at a stationary point x (discussed below), there exist $\tilde{\mu} = (\tilde{\mu}_1, \tilde{\mu}_2, \ldots, \tilde{\mu}_{m+r}) \in \mathcal{R}^{m+r}$ that satisfy the Karush-Kuhn-Tucker (KKT) conditions [3, 4, 15]:

$$\begin{array}{c} \nabla_x \tilde{f}(x) + \sum_{i=1}^{m} \tilde{\mu}_i \nabla_x \tilde{g}_i(x) + \sum_{j=1}^{r} \tilde{\mu}_{m+j} \nabla_x \tilde{h}_j(x) = 0, \ \ \tilde{h}(x) = 0; \\ \tilde{\mu}_i \geq 0, \ \ \tilde{g}_i(x) \leq 0, \ \ \tilde{\mu}_i \tilde{g}_i(x) = 0, \quad \text{for } i = 1, 2, \ldots, m. \end{array} \tag{1.8}$$

In this case, $\tilde{\mu}$ are referred to as the Lagrange multipliers, and x is called a Karush-Kuhn-Tucker (KKT) point. We denote the set of Lagrange multipliers by

$$\Lambda(x) = \left\{ \tilde{\mu} \in \mathcal{R}^{m+r} \mid \ \tilde{\mu} \text{ satisfies (1.8) at } x \ \right\}. \tag{1.9}$$

A simple inspection of the definitions of $\Lambda(x)$ and $\Lambda_1^g(x)$ reveals that:

$$\tilde{\mu} \in \Lambda(x) \Leftrightarrow (1, \tilde{\mu}) \in \Lambda_1^g(x).$$

Also, because of the first-order homogeneity of the conditions (1.2), and from (1.8), it immediately follows that

$$\Lambda(x) \neq \emptyset \Leftrightarrow \Lambda_1^g(x) \neq \emptyset \Leftrightarrow \exists \tilde{\lambda} \in \Lambda^g(x), \text{ such that } \tilde{\lambda}_0 \neq 0. \tag{1.10}$$

The regularity condition, or constraint qualification, ensures that a linear approximation of the feasible set in the neighborhood of a stationary point x captures the geometry of the feasible set. The regularity condition that we will use at times at a stationary point x is the Mangasarian-Fromovitz constraint qualification (MFCQ) [33, 32]:

$$\text{(MFCQ)} \qquad \begin{array}{l} 1. \ \nabla_x \tilde{h}_j(x), \ j = 1, 2, \ldots, r, \ \text{are linearly independent and} \\ 2. \ \exists p \neq 0 \ \text{ such that } \ \nabla_x \tilde{h}_j(x)^T p = 0, \ \ j = 1, 2, \ldots, r \\ \text{and } \ \nabla_x \tilde{g}_i(x)^T p < 0, \ \ i \in \tilde{\mathcal{A}}(x). \end{array}$$

It is well known [23] that (MFCQ) is equivalent to the fact that the set $\Lambda(x)$ of Lagrange multipliers of (1.1) is nonempty and bounded at a stationary point x of (1.1). Note that $\Lambda(x)$ is certainly polyhedral in any case.

Another condition that we will use on occasion is the strict Mangasarian-Fromovitz constraint qualification (SMFCQ). We say that this condition is satisfied by (1.1) at a KKT point x if

$$\text{(SMFCQ)} \qquad \begin{array}{l} \text{1) (MFCQ) is satisfied at } x \text{ and} \\ \text{2) the Lagrange multiplier set } \Lambda(x) \text{ contains exactly one element.} \end{array}$$

The critical cone at a stationary point x is [14, 42]

$$\begin{aligned} \mathcal{C}(x) \;=\; & \Big\{ u \in \mathcal{R}^n \mid \nabla_x \tilde{h}_j(x)^T u = 0,\ j = 1,2,\ldots,r, \\ & \nabla_x \tilde{g}_i(x)^T u \le 0,\ i \in \tilde{\mathcal{A}}(x);\ \nabla_x \tilde{f}(x)^T u \le 0 \Big\}. \end{aligned} \tag{1.11}$$

We now review the conditions for a point x^* to be a solution of (1.1). The second-order necessary conditions for x^* to be a local minimum are that $\Lambda^g(x^*) \neq \emptyset$ and [26]

$$\forall u \in \mathcal{C}(x^*),\ \ \exists \tilde{\lambda}^* \in \Lambda^g(x^*),\ \text{such that}\ u^T \nabla^2_{xx} \mathcal{L}(x^*, \tilde{\lambda}^*) u \ge 0. \tag{1.12}$$

The second-order sufficient conditions for x^* to be a local minimum are that $\Lambda^g(x^*) \neq \emptyset$ and [26]

$$\forall u \in \mathcal{C}(x^*),\ \ u \neq 0,\ \ \exists \tilde{\lambda}^* \in \Lambda^g(x^*),\ \text{such that}\ u^T \nabla^2_{xx} \mathcal{L}(x^*, \tilde{\lambda}^*) u > 0. \tag{1.13}$$

Stronger second-order conditions are Robinson's conditions. These conditions are that, at a solution x^*, the following condition holds:

$$\forall u \in \mathcal{C}(x^*),\ \ u \neq 0,\ \ \forall \tilde{\lambda}^* \in \Lambda^g(x^*),\ \text{we have that}\ u^T \nabla^2_{xx} \mathcal{L}(x^*, \tilde{\lambda}^*) u > 0.$$

In a fact we will invoke Robinson's conditions for the case where $\Lambda^g_1(x^*) \neq \emptyset$. In the latter situation, Robinson's conditions are equivalent to:

$$\text{(RSOSC)} \quad \forall u \in \mathcal{C}(x^*),\ \ u \neq 0,\ \ \forall \tilde{\lambda}^* \in \Lambda^g_1(x^*),\ \text{we have that}\ u^T \nabla^2_{xx} \mathcal{L}(x^*, \tilde{\lambda}^*) u > 0.$$

1.2. Notation. For a mapping $q : \mathcal{R}^n \to \mathcal{R}^l$, we define:

$$q^+(x) = \begin{pmatrix} \max\{q_1(x), 0\} \\ \max\{q_2(x), 0\} \\ \vdots \\ \max\{q_l(x), 0\} \end{pmatrix} \text{ and } if q^-(x) = \begin{pmatrix} \max\{-q_1(x), 0\} \\ \max\{-q_2(x), 0\} \\ \vdots \\ \max\{-q_l(x), 0\} \end{pmatrix}.$$

With this definition, it immediately follows that $q(x) = q^+(x) - q^-(x)$ and that $|q_i(x)| = q_i^+(x) + q_i^-(x)$, $i = 1, 2, \ldots, l$.

In this work we use the convention that $\nabla_x q(x)$ is a matrix with l rows and n columns. This will allow us to write first order Taylor expansion of $q(x)$ at some point $\hat{x}$ as $q(\hat{x}) + \nabla_x q(\hat{x})\,(x - \hat{x})$, without the need to use a transpose sign. In particular, if $q : \mathcal{R}^n \to \mathcal{R}$, $\nabla_x q(x)$ is a row vector.

We will also use the Landau notation. We say that a is of order b which we denote by $a = O(b)$, if there exists c such that $a \leq cb$ for all a and b sufficiently small. We denote by $a = \Omega(b)$ quantities a and b that satisfy $a = O(b)$ and $b = O(a)$.

We will use certain symbols twice, to denote related data of different programs. However, to avoid confusion, we will use a ˜ sign for the data of the general nonlinear programming problem (1.1), whereas the same objects associated with the (MPCC) problem (to be defined later) are denoted it without the ˜ sign. For instance, $\tilde{f},\ \tilde{g},\ \tilde{h}$, denote, respectively, the objective, the inequality constraint, and the equality constraints of the general nonlinear programming problem, whereas f, g, h denote, respectively, the objective, the inequality constraints, and the equality constraints of the MPCC problem. The MPCC problem, however, has, in addition, complementarity constraints.

In the literature problems of the type we treat here are also called mathematical programs with equilibrium constraints (MPEC), an acronym that we may use when we invoke optimality conditions from the respective references.

We denote the L_∞ nondifferentiable penalty function by

$$\tilde{P}_\infty(x) = \max\left\{\tilde{g}_1(x), \tilde{g}_2(x), ..., \tilde{g}_m(x), \left|\tilde{h}_1(x)\right|, \left|\tilde{h}_2(x)\right|, \ldots, \left|\tilde{h}_r(x)\right|, 0\right\}. \tag{1.14}$$

We also define the L_1 penalty function as

$$\tilde{P}_1(x) = \sum_{i=1}^{m} \tilde{g}_i^+(x) + \sum_{j=1}^{r} \left|\tilde{h}_j(x)\right|. \tag{1.15}$$

It is immediate that:

$$0 \leq \tilde{P}_\infty(x) \leq \tilde{P}_1(x) \leq (m+r)\tilde{P}_\infty(x).$$

An obvious consequence of (1.15) and (1.14) is that x is a feasible point of (1.1) if and only if $\tilde{P}_1(x) = \tilde{P}_\infty(x) = 0$.

We say that the nonlinear program (1.1) satisfies the quadratic growth condition with a parameter $\tilde{\sigma}$ at x^* if

$$\max\left\{\tilde{f}(x) - \tilde{f}(x^*), \tilde{P}_\infty(x)\right\} \geq \tilde{\sigma}\, ||x - x^*||^2 \tag{1.16}$$

holds for some $\tilde{\sigma} > 0$ and all x in a neighborhood of x^*. The quadratic growth condition is equivalent to the second-order sufficient conditions (1.13), [6, 7, 26, 27, 42] and it is the weakest possible second-order condition.

For the case in which (MFCQ) holds at a solution x^* of (1.1), the quadratic growth condition at x^* is equivalent to [6]

$$\tilde{f}(x) - \tilde{f}(x^*) \geq \tilde{\sigma}_{\tilde{f}}\, ||x - x^*||^2 \tag{1.17}$$

for some $\tilde{\sigma}_{\tilde{f}} > 0$ and all x feasible in a neighborhood of x^*.

1.3. Exact Penalty Conditions for Degenerate Nonlinear Programming. We now assume that at a solution x^* of the nonlinear program (1.1) the following conditions hold:

1. The Lagrange multiplier set at x^*, $\Lambda(x^*)$, is not empty.
2. The quadratic growth condition (1.16) is satisfied.

Then there exists a neighborhood $\mathcal{V}(x^*)$, some penalty parameters $\tilde{c}_1 \geq 0$, $\tilde{c}_\infty \geq 0$ and some growth parameters $\sigma_1 > 0$ and $\sigma_\infty > 0$ such that [7, Theorem 3.113]

$$\forall x \in \mathcal{V}(x^*), \psi_1(x) = \tilde{f}(x) + \tilde{c}_1 \tilde{P}_1(x) \geq \tilde{f}(x^*) + \sigma_1 \|x - x^*\|^2 \tag{1.18}$$
$$= \psi_1(x^*) + \sigma_1 \|x - x^*\|^2,$$
$$\forall x \in \mathcal{V}(x^*), \psi_\infty(x) = \tilde{f}(x) + \tilde{c}_\infty \tilde{P}_\infty(x) \geq \tilde{f}(x^*) + \sigma_\infty \|x - x^*\|^2 \tag{1.19}$$
$$= \psi_\infty(x^*) + \sigma_\infty \|x - x^*\|^2.$$

Therefore, x^* becomes an unconstrained strict local minimum for the nondifferentiable functions $\psi_1(x)$ and $\psi_\infty(x)$. Such functions are called nondifferentiable exact merit functions for the nonlinear program (1.1) [3, 4, 15]. If (1.18) and (1.19) are satisfied then we say that the functions $\psi_1(x)$ and $\psi_\infty(x)$ satisfy a quadratic growth condition near x^*.

1.4. Formulation of Mathematical Programs with Complementarity Constraints. We use notation similar to the one in [40] to define a mathematical program with complementarity constraints (MPCC).

$$\begin{array}{llll}
(\text{MPCC}) \quad \min_x & f(x) & & \\
\text{subject to} & g_i(x) & \leq 0, & i = 1, 2, \ldots, n_i \\
& h_j(x) & = 0, & j = 1, 2, \ldots, n_e \\
& F_{k,1}(x) & \leq 0, & k = 1, 2, \ldots, n_c \\
& F_{k,2}(x) & \leq 0, & k = 1, 2, \ldots, n_c \\
& F_{k,1}(x) F_{k,2}(x) & \leq 0, & k = 1, 2, \ldots, n_c.
\end{array}$$

In this work we assume that the data of (MPCC) ($f(x), h(x), g(x)$ and $F_{k,i}(x)$, for $k = 1, 2, \ldots, n_c$, and $i = 1, 2$) are twice continuously differentiable.

For a given k, the constraints $F_{k,1}(x) \leq 0$, $F_{k,2}(x) \leq 0$ imply that $F_{k,1}(x)F_{k,2}(x) \leq 0$ is equivalent to $F_{k,1}(x)F_{k,2}(x) = 0$. The constraints $F_{k,1}(x)F_{k,2}(x) \leq 0$ are thus called complementarity constraints and are active at any feasible point of (MPCC). Therefore (MPCC) included a particular choice of representing the complementarity constraints $F_{k1}(x) \perp F_{k2}(x)$ for $k = 1, 2, \ldots, n_c$ as constraints of a smooth nonlinear program. We present another one of the several equivalent smooth nonlinear programming formulations in (2.13).

Since we cannot have $F_{k,1}(x) < 0$, $F_{k,2}(x) < 0$, and $F_{k,1}(x)F_{k,2}(x) < 0$ simultaneously, it follows that (MFCQ) cannot hold at any feasible point x [30, 40].

1.5. MPCC Notation. In this section, which previews our general convergence results, we use the same notation from [40] to denote certain index sets, because at some point we invoke a theorem from that reference. Later, in our super linear convergence results we will use notation from [18] to denote similar index sets, because we will use results from the latter reference.

If i is one of $1, 2$ we define $\bar{i} = 2 - i + 1$. Therefore $i = 1 \Rightarrow \bar{i} = 2$, and $i = 2 \Rightarrow \bar{i} = 1$. The complementarity constraints can thus be written as $F_{k,i}(x)F_{k,\bar{i}}(x) \leq 0$, $k = 1, 2, \ldots, n_c$. We use the notation

$$F(x) = \left(F_{11}(x), F_{12}(x), F_{21}(x), F_{22}(x), \ldots, F_{n_c1}(x), F_{n_c2}(x)\right)^T. \tag{1.20}$$

The active set of the inequality constraints $g_i(x) \leq 0$, $1 \leq i \leq m$, at a feasible point x is

$$\mathcal{A}(x) = \{i \in \{1, 2, \ldots, n_i\} \mid g_i(x) = 0\}. \tag{1.21}$$

We use the following notation:

$$(1.22)\quad \mathcal{I}(x) = \left\{(k,i) \in \{1,2,\ldots,n_c\} \times \{1,2\} \mid F_{k,\bar{i}}(x) < 0\right\},$$

$$(1.23)\quad \overline{\mathcal{I}}(x) = \{(k,i) \in \{1,2,\ldots,n_c\} \times \{1,2\} \mid F_{k,i}(x) < 0\},$$

$$(1.24)\quad \overline{\mathcal{D}}(x) = \left\{(k,i) \in \{1,2,\ldots,n_c\} \times \{1,2\} \mid F_{k,i}(x) = F_{k,\bar{i}}(x) = 0\right\},$$

$$(1.25)\quad \mathcal{I}^c(x) = \{1,2,\ldots,n_c\} \times \{1,2\} - \mathcal{I}(x),$$

$$(1.26)\quad \mathcal{K}(x) = \{k \in \{1,2,\ldots,n_c\} \mid (k,1) \in \mathcal{I}(x) \ \text{ or } \ (k,2) \in \mathcal{I}(x)\},$$

$$(1.27)\quad \overline{\mathcal{K}}(x) = \{k \in \{1,2,\ldots,n_c\} \mid F_{k,1}(x) = F_{k,2}(x) = 0\} = \{1,2,\ldots,n_c\} - \mathcal{K}(x).$$

There are two cases for the constraints involved in the complementarity constraints at a feasible point x.

1. $F_{k,1}(x)+F_{k,2}(x) < 0$. In this case there is an $i(k) \in \{1,2\}$ such that $F_{k,i(k)} = 0$ and $F_{k,\bar{i}(k)} < 0$. Therefore, with our notation $k \in \mathcal{K}(x)$, $(k, i(k)) \in \mathcal{I}(x)$ and $(k, \bar{i}(k)) \in \overline{\mathcal{I}}(x)$. We call $F_{k,1}(x), F_{k,2}(x)$ a nondegenerate (or strictly complementary) pair. In the rest of the paper $i(k)$ and $\bar{i}(k)$ will have the meaning defined in this paragraph, whenever $k \in \mathcal{K}$.
2. $F_{k,1}(x) + F_{k,2}(x) = 0$, or $F_{k,1}(x) = F_{k,2}(x) = 0$. In this case $k \in \overline{\mathcal{K}}(x)$, $(k,1) \in \overline{\mathcal{D}}(x)$ and $(k,2) \in \overline{\mathcal{D}}(x)$. We call $F_{k,1}(x), F_{k,2}(x)$ a degenerate pair.

Therefore $\mathcal{I}(x)$ and $\overline{\mathcal{I}}(x)$ contain the indices of the active constraints at which strict complementarity occurs, whereas $\overline{\mathcal{D}}(x)$ contains the indices of the constraints that are degenerate at x from the point of view of complementarity. The set $\mathcal{K}(x)$ represents the indices k at which strict complementarity occurs and $\overline{\mathcal{K}}(x)$ the indices k at which complementarity degeneracy occurs.

Since we are interested in the behavior of (MPCC) at a solution point x^*, we may avoid the dependence of these index sets on x. Therefore we denote $\mathcal{I} = \mathcal{I}(x^*)$, $\overline{\mathcal{D}} = \overline{\mathcal{D}}(x^*)$, $\mathcal{K} = \mathcal{K}(x^*)$, and $\mathcal{A} = \mathcal{A}(x^*)$.

For a set of pairs $\mathcal{J} \subset \{1,2,\ldots,n_c\} \times \{1,2\}$ we denote by $F_{\mathcal{J}}$ a map whose components are $F_{k,i}$ with $(k,i) \in \mathcal{J}$.

1.6. Associated Nonlinear Programs at x^*. In this section we associate two nonlinear programs to (MPCC). This will help with characterizing the stationarity conditions for (MPCC). The notation is from [40].

At x^* we associate the relaxed nonlinear program (RNLP) to (MPCC).

$$\begin{array}{llll}
\text{(RNLP)} \qquad & \min_x & f(x) & \\
& \text{subject to} & g_i(x) \le 0, & i = 1,2,\ldots,n_i \\
& & h_j(x) = 0, & j = 1,2,\ldots,n_e \\
& & F_{\overline{\mathcal{D}}}(x) \le 0, & \\
& & F_{\mathcal{I}}(x) = 0. &
\end{array}$$

As it can be seen, (RNLP) is obtained from (MPCC) by dropping the elements from $F(x)$ that are inactive at x^*, as well as the complementarity constraints, but enforcing the complements of inactive constraints as equality constraints.

Note that (RNLP) typically includes the constraints $F_{\bar{I}}(x) \le 0$ in order to have a tighter relaxation of the feasible set of (MPCC). The nonlinear program (RNLP), however, is used only for local analysis near a solution x^*. Since the constraints $F_{\bar{I}}(x) \le 0$ are not active in a neighborhood of x^*, we do not consider them in the analysis.

We also associate at x^* the tightened nonlinear program (TNLP), in which all the complementarity constraints in (MPCC) are dropped and all active constraints at x^* connected to complementarity constraints are replaced by equality constraints.

$$
\text{(TNLP)} \qquad \begin{array}{rlll} \min_x & f(x) & & \\ \text{subject to} & g_i(x) & \leq 0, & i = 1, 2, \ldots, n_i \\ & h_j(x) & = 0, & j = 1, 2, \ldots, n_e \\ & F_{\overline{\mathcal{D}}}(x) & = 0, & \\ & F_{\mathcal{I}}(x) & = 0. & \end{array}
$$

We immediately see that, near x^*, (TNLP) is a more constrained problem than (MPCC), which in turn is more constrained than (RNLP), and all three programs have the same objective function. As a result, if x^* is a local solution of (RNLP), then it must be a local solution of (MPCC). Also, if x^* is a local solution of (MPCC), then it will be a local solution of (TNLP). None of the reverse implications hold in general for either local solutions or stationary points.

However, if (TNLP) satisfies (SMFCQ) at a solution x^* of (MPCC), then x^* is a Karush-Kuhn-Tucker point of (TNLP) and (RNLP) [40].

2. The Lagrange Multiplier Set of (MPCC). In this section we analyze the relationship between the relevant mathematical objects of (MPCC) and (RNLP) at a solution x^*. The (RNLP) formulation does not immediately violate (MFCQ), the way (MPCC) does. By establishing a correspondence between the Lagrange multiplier sets of (RNLP) and (MPCC) we ensure that, under certain conditions, (MPCC) has a nonempty Lagrange multiplier set, although it does not satisfy a constraint qualification.

2.1. Critical Cones. In this section we compare the critical cones of (MPCC) and (RNLP). The active sets play a structural part in the definition of the critical cones. We have that:

$$
\nabla_x \left(F_{k,1} F_{k,2} \right) (x^*) = F_{k,1}(x^*) \nabla_x F_{k,2}(x^*) + F_{k,2}(x^*) \nabla_x F_{k,1}(x^*).
$$

Using the definition (1.11) we get that the critical cone of (MPCC) is

$$
\mathcal{C}_{\mathsf{MPCC}} = \{ u \in R^n \mid \begin{array}{lcll} \nabla_x f(x^*) u & \leq & 0, & \\ \nabla_x g_i(x^*) u & \leq & 0, & i \in \mathcal{A} \\ \nabla_x h_j(x^*) u & = & 0, & j \in 1, 2, \ldots, n_e \\ \nabla_x F_{k,1}(x^*) u & \leq & 0, & k \in \overline{\mathcal{K}} \\ \nabla_x F_{k,2}(x^*) u & \leq & 0, & k \in \overline{\mathcal{K}} \\ \nabla_x F_{k,i(k)}(x^*) u & \leq & 0, & (k, i(k)) \in \mathcal{I} \\ F_{k,\bar{i}(k)}(x^*) \nabla_x F_{k,i(k)}(x^*) u & \leq & 0, & (k, i(k)) \in \mathcal{I} \} . \end{array} \tag{2.1}
$$

Note that the definition that we use here of the critical cone corresponds to the nonlinear programming interpretation of (MPCC). There exists another, combinatorial, definition of the critical cone, that is used in connection to disjunctive approaches [30, Equation 5.3.(2)].

We use (1.11) again to determine the critical cone of the relaxed nonlinear program. It is immediate from the definition of the index sets $\mathcal{I}, \mathcal{K}$, and $\overline{\mathcal{D}}$ that all constraints involving components of $F(x)$ are active at x^* for (RNLP). It thus follows that the critical cone of (RNLP) is

$$
\mathcal{C}_{\mathsf{RNLP}} = \{u \in R^n \mid \begin{array}{lcll} \nabla_x f(x^*)u & \leq & 0, & \\ \nabla_x g_i(x^*)u & \leq & 0, & i \in \mathcal{A} \\ \nabla_x h_j(x^*)u & = & 0, & j \in 1,2,\ldots,n_e \\ \nabla_x F_{k,1}(x^*)u & \leq & 0, & k \in \overline{\mathcal{K}} \\ \nabla_x F_{k,2}(x^*)u & \leq & 0, & k \in \overline{\mathcal{K}} \\ \nabla_x F_{k,i(k)}(x^*)u & = & 0, & (k,i(k)) \in \mathcal{I}\}. \end{array} \tag{2.2}
$$

Lemma 2.1. $\mathcal{C}_{\mathsf{MPCC}} = \mathcal{C}_{\mathsf{RNLP}}$.

Proof The conclusion is immediate, by noting that all the constraints involving the critical cones are the same with the exception of the ones involving indices k for which $(k,i(k)) \in \mathcal{I}$. For these k, from the definition (1.22) of the index sets it follows that $F_{k,\bar{i}(k)}(x^*) < 0$. We therefore have that:

$$
\begin{array}{lclcrcl} \nabla_x F_{k,i(k)}(x^*)u & \leq & 0 & \text{and} & F_{k,\bar{i}(k)}(x^*)\nabla_x F_{k,i(k)}(x^*)u & \leq & 0 \Leftrightarrow \\ \nabla_x F_{k,i(k)}(x^*)u & \leq & 0 & \text{and} & \nabla_x F_{k,i(k)}(x^*)u & \geq & 0 \Leftrightarrow \\ & & & & \nabla_x F_{k,i(k)}(x^*)u & = & 0. \end{array}
$$

Since the remaining constraints of (RNLP) and (MPCC) are the same this equivalence proves the claim. $\diamond$

2.2. Generalized Lagrange Multipliers. The set of generalized Lagrange multipliers of (MPCC) at x^* is a set of multiples

$$
0 \neq (\alpha, \nu, \pi, \mu, \eta) \in \mathcal{R} \times \mathcal{R}^{n_i} \times \mathcal{R}^{n_e} \times \mathcal{R}^{2n_c} \times \mathcal{R}^{n_c}
$$

that satisfies the Fritz-John conditions (1.2). Since μ are the multipliers corresponding to the components of $F(x)$, we will index them by elements in $(1,2,\ldots,n_c) \times (1,2)$. The Fritz-John conditions for (MPCC) at x^* are that x^* is feasible for (MPCC) and that

$$
\alpha \nabla_x f(x^*) + \sum_{i=1}^{n_i} \nu_i \nabla_x g_i(x^*) + \sum_{j=1}^{n_e} \pi_j \nabla_x h_j(x^*) + \tag{2.3}
$$

$$
\sum_{k=1}^{n_c} \left[\mu_{k,1}\nabla_x F_{k,1}(x^*) + \mu_{k,2}\nabla_x F_{k,2}(x^*) + \eta_k \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*)\right] = 0
$$

$$
\begin{array}{rrrl} F_{k,i}(x^*) \leq 0, & \mu_{k,i} \geq 0, & \mu_{k,i}F_{k,i}(x^*) = 0, & k = 1,2,\ldots,n_c, \\ & & & i = 1,2 \\ g_i(x^*) \leq 0, & \nu_i \geq 0, & \nu_i g_i(x^*) = 0, & i = 1,2,\ldots,n_i \\ F_{k,1}(x^*)F_{k,2}(x^*) \leq 0, & \eta_k \geq 0, & \eta_k F_{k,1}(x^*)F_{k,2}(x^*) = 0, & k = 1,2,\ldots,n_c. \end{array} \tag{2.4}
$$

From our definition of the index sets it follows that $F_{\overline{\mathcal{I}}}(x^*) < 0$ and $g_{\mathcal{A}^c}(x^*) < 0$. Therefore, from the complementarity conditions (2.4), it follows that $\mu_{\overline{\mathcal{I}}} = 0$ and $\nu_{\mathcal{A}^c} = 0$.

We can also determine the relations satisfied by the generalized Lagrange multipliers of (RNLP). As discussed above, the index sets that define (RNLP) have been chosen such that all constraints involving components of $F(x)$ are active. Therefore the generalized Lagrange multipliers are:

$$
0 \neq (\tilde{\alpha}, \tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta}) \in \mathcal{R} \times \mathcal{R}^{n_i} \times \mathcal{R}^{n_e} \times \mathcal{R}^{n_{\overline{\mathcal{D}}}} \times \mathcal{R}^{n_{\mathcal{I}}}
$$

that satisfy the Fritz-John conditions:

$$\tilde{\alpha}\nabla_x f(x^*) + \sum_{i=1}^{n_i} \tilde{\nu}_i \nabla_x g_i(x^*) + \sum_{j=1}^{n_e} \tilde{\pi}_j \nabla_x h_j(x^*) + \sum_{k\in\overline{\mathcal{K}}} \left[\tilde{\mu}_{k,1}\nabla_x F_{k,1}(x^*) + \tilde{\mu}_{k,2}\nabla_x F_{k,2}(x^*)\right] + \sum_{k\in\mathcal{K}} \tilde{\eta}_{k,i(k)}\nabla_x F_{k,i(k)}(x^*) = 0 \tag{2.5}$$

$$\begin{array}{llll} g_i(x^*) \leq 0, & \tilde{\nu}_i \geq 0, & \tilde{\nu}_i g_i(x^*) = 0, & i = 1, 2, \ldots, n_i \\ & \tilde{\mu}_{k,1} \geq 0, & \tilde{\mu}_{k,2} \geq 0, & k \in \overline{\mathcal{K}}. \end{array} \tag{2.6}$$

Here $\tilde{\mu}$ is a vector that is indexed by elements of $\overline{\mathcal{D}}$, and $\tilde{\eta}$ is indexed by elements of $\mathcal{I}$.

2.3. Relations between the generalized Lagrange Multiplier Sets of (MPCC) and (RNLP). Take $\tilde{\lambda} = (\tilde{\alpha}, \tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta}) \in \Lambda^g{}_{\mathsf{RNLP}}$. We construct from the generalized multiplier $\tilde{\lambda}$ of (RNLP) a generalized multiplier λ° of (MPCC). We define the following types of components of λ°.

1. Components that correspond to the objective function or the inequality constraints $g_i(x) \leq 0$ and equality constraints $h_j(x) = 0$

$$\alpha^\circ = \tilde{\alpha}; \;\; \nu^\circ = \tilde{\nu}; \;\; \pi^\circ = \tilde{\pi}. \tag{2.7}$$

2. Components connected to the pairwise degenerate constraints. For these we have $k \in \overline{\mathcal{K}}$ and $(k,1),(k,2) \in \overline{\mathcal{D}}$ or $F_{k,1}(x^*) = F_{k,2}(x^*) = 0$. We define

$$\mu^\circ_{k,i} = \tilde{\mu}_{k,i}, \; (k,i) \in \overline{\mathcal{D}}; \quad \eta^\circ_k = 0, \; k \in \overline{\mathcal{K}}. \tag{2.8}$$

We have that:

$$\nabla_x \left(F_{k,1}F_{k,2}\right)(x^*) = 0,$$

and therefore

$$\begin{array}{r} \tilde{\mu}_{k,1}\nabla_x F_{k,1}(x^*) + \tilde{\mu}_{k,2}\nabla_x F_{k,2}(x^*) = \mu^\circ_{k,1}\nabla_x F_{k,1}(x^*) \quad + \\ \mu^\circ_{k,2}\nabla_x F_{k,2}(x^*) + \eta^\circ_k \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*). \end{array} \tag{2.9}$$

3. Components connected to pairwise strictly complementary constraints. In this case we have $k \in \mathcal{K}$, $(k,i(k)) \in \mathcal{I}$, and $(k,\bar{i}(k)) \in \overline{\mathcal{I}}$. Therefore $F_{k,\bar{i}(k)}(x^*) < 0$, $F_{k,i(k)}(x^*) = 0$, and we thus define the multipliers

$$\begin{array}{rcll} \mu^\circ_{k,i(k)} & = & \max\left\{\tilde{\eta}_{k,i(k)}, 0\right\}, & (k,i(k)) \in \mathcal{I} \\ \mu^\circ_{k,\bar{i}(k)} & = & 0, & (k,\bar{i}(k)) \in \overline{\mathcal{I}} \\ \eta^\circ_k & = & \frac{1}{F_{k,\bar{i}(k)}(x^*)} \min\left\{\tilde{\eta}_{k,i(k)}, 0\right\}, & k \in \mathcal{K}. \end{array} \tag{2.10}$$

It is immediate from these definitions that $\mu^\circ_{k,i(k)} \geq 0$ and $\eta^\circ_k \geq 0$. Since, for fixed k, $\tilde{\eta}_{k,i(k)}$ is the only multiplier of (RNLP) involved in definition (2.10), we obtain that

$$\begin{array}{rl} \tilde{\eta}_{k,i(k)}\nabla_x F_{k,i(k)}(x^*) = & \left[\max\left\{\tilde{\eta}_{k,i(k)}, 0\right\} + \min\left\{\tilde{\eta}_{k,i(k)}, 0\right\}\right] \nabla_x F_{k,i(k)}(x^*) \\ = & \mu^\circ_{k,i(k)}\nabla_x F_{k,i(k)}(x^*) + \eta^\circ_k F_{k,\bar{i}(k)}(x^*)\nabla_x F_{k,i(k)}(x^*) \quad = \\ \multicolumn{2}{l}{\mu^\circ_{k,i(k)}\nabla_x F_{k,i(k)}(x^*) + \mu^\circ_{k,\bar{i}(k)}\nabla_x F_{k,\bar{i}(k)}(x^*) + \eta^\circ_k \nabla_x \left(F_{k,i(k)}F_{k,\bar{i}(k)}\right)(x^*).} \end{array} \tag{2.11}$$

After we compare the terms that, following (2.9) and (2.11), are equal in (2.5) and (2.3), we get that $\lambda^\circ = (\alpha^\circ, \nu^\circ, \pi^\circ, \mu^\circ, \eta^\circ)$ satisfies (2.3) as well as (2.4). By tracing the definition of λ° we also have that $\tilde{\lambda} \neq 0 \Rightarrow \lambda^\circ \neq 0$. Therefore λ° is a generalized Lagrange multiplier of (MPCC) or

$$\lambda^\circ = (\alpha^\circ, \nu^\circ, \pi^\circ, \mu^\circ, \eta^\circ) \in \Lambda^g_{\mathsf{MPCC}}, \tag{2.12}$$

where $\alpha^\circ = \tilde{\alpha}$ from (2.7).

THEOREM 2.2. *The set of Lagrange multipliers of (RNLP) is not empty if and only if the set of Lagrange multipliers of (MPCC) is not empty.*

Proof If the Lagrange multiplier set of (RNLP) is not empty, we can choose $\tilde{\lambda} = (1, \tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta}) \in \Lambda^g_{1,\ \mathsf{RNLP}}$. From (2.12) it follows that $\lambda^\circ = (1, \nu^\circ, \pi^\circ, \mu^\circ, \eta^\circ) \in \Lambda^g_{1,\mathsf{MPCC}}$ is a generalized multiplier of (MPCC). From (1.10) it follows that the Lagrange multiplier set of (MPCC) is not empty. The proof of the reverse statement follows in a similar manner and is omitted. ◇

One one hand, in our work, we are interested in KKT points of (MPCC). On the other, existence of Lagrange multipliers for (RNLP) at x^* amounts to what is called strong stationarity in the literature on MPCCs [40]. Therefore our previous result says that x^* is stationary for (MPCC) if and only if it is strongly stationary.

COROLLARY 2.3. *Assume that (TNLP) satisfies (SMFCQ) at a solution x^* of (MPCC), i.e.*

1. $\nabla_x F_{\overline{\mathcal{D}}}(x^*)$, $\nabla_x F_{\mathcal{I}}(x^*)$, *and* $\nabla_x h(x^*)$ *are linearly independent.*
2. *There exists* $p \neq 0$ *such that* $\nabla_x F_{\overline{\mathcal{D}}}(x^*) p = 0$, $\nabla_x F_{\mathcal{I}}(x^*) p = 0$, $\nabla_x h(x^*) p = 0$, $\nabla_x g_i(x^*) p < 0$, *for* $i \in \mathcal{A}(x^*)$.
3. *The Lagrange multiplier set of (TNLP) at x^* has a unique element.*

Then the Lagrange multiplier set of (MPCC) is not empty.

Proof From [40, Theorem 2], since (TNLP) satisfies (SMFCQ) at x^*, the Lagrange multiplier set of (RNLP) is not empty. Following Theorem 2.2, we obtain that the Lagrange multiplier set of (MPCC) is not empty, which proves the claim. ◇

The proof of the Corollary 2.3 also follows immediately from [40, Theorem 4] after using the observation that, when (SMFCQ) holds, x^* is a B-stationary point of (MPCC), [40, Section 2.1]. The conclusion of Corollary 2.3 does not hold in absence of (SMFCQ) [40, Example 3].

2.4. An alternative formulation. We also investigate the following equivalent formulation of (MPCC), where the complementarity constraints have been replaced by one constraint:

$$\begin{array}{lrll} \min_x & f(x) & & \\ \text{subject to} & g_i(x) & \leq 0, & i = 1, 2, \ldots, n_i \\ & h_j(x) & = 0, & j = 1, 2, \ldots, n_e \\ & F_{k,1}(x) & \leq 0, & k = 1, 2, \ldots, n_c \\ & F_{k,2}(x) & \leq 0, & k = 1, 2, \ldots, n_c \\ & \sum_{k=1}^{n_c} F_{k,1}(x) F_{k,2}(x) & \leq 0. & \end{array} \tag{2.13}$$

At a feasible point of the above program, we must have that $\sum_{k=1}^{n_c} F_{k,1}(x) F_{k,2}(x) = 0$ and the equivalence between (2.13) and (MPCC) follows immediately. This formulation is of interest in computations because it has less constraints than (MPCC).

LEMMA 2.4. *If the Lagrange multiplier set of (MPCC) is not empty, there exists a generalized Lagrange multiplier* $(1, \nu, \pi, \mu, \eta) \in \Lambda^g_{\mathsf{MPCC}}$ *such that* $\eta_k = \eta_1$, $k = 2, 3, \ldots, n_c$.

Proof The proof follows from straightforward algebra of the type used in Subsection 2.2. $\diamond$

We now describe the generalized Lagrange multiplier set of the alternative formulation (2.13). We denote mathematical objects connected to (2.13) by the subscript $_{MPCC1}$. We write the Fritz-John conditions (1.2) for (2.13) at the point x^*, and we obtain

$$\begin{aligned}(2.14)\qquad & \alpha^\diamond \nabla_x f(x^*) + \sum_{i=1}^{n_i} \nu_i^\diamond \nabla_x g_i(x^*) + \sum_{j=1}^{n_e} \pi_j^\diamond \nabla_x h_j(x^*) + \\ & \sum_{k=1}^{n_c} \sum_{i=1}^{2} \mu_{k,i}^\diamond \nabla_x F_{k,i}(x^*) + \eta_1^\diamond \sum_{k=1}^{n_c} \nabla_x \left(F_{k,1}F_{k,2}\right)(x^*) = 0\end{aligned}$$

$$(2.15)\qquad \begin{array}{llll} F_{k,i}(x^*) \leq 0, & \mu_{k,i}^\diamond \geq 0, & \mu_{k,i}^\diamond F_{k,i}(x^*) = 0, & k = 1,2,\ldots,n_c, \\ & & & i = 1,2 \\ g_i(x^*) \leq 0, & \nu_i^\diamond \geq 0, & \nu_i^\diamond g_i(x^*) = 0, & i = 1,2,\ldots,n_i \end{array}$$

and $\eta_1^\diamond \geq 0 \in \mathcal{R}$. Note that from the fact that x^* is feasible for (2.13) we also have that $\sum_{k=1}^{n_c} F_{k,1}(x)F_{k,2}(x) \leq 0$ and $F_{ki} \leq 0$, for $k = 1,2,\ldots,n_c$ and $i = 1,2$. Consequently, we have that $\sum_{k=1}^{n_c} F_{k,1}(x)F_{k,2}(x) = 0$ and thus the complementarity condition $\eta_1^\diamond \sum_{k=1}^{n_c} F_{k,1}(x)F_{k,2}(x) = 0$ is redundant. We therefore ignore it in the Fritz-John conditions (2.14–2.15).

A generalized multiplier of (2.13) is thus

$$\lambda^\diamond = (\alpha^\diamond, \nu^\diamond, \pi^\diamond, \mu^\diamond, \eta_1^\diamond) \in \Lambda^g_{MPCC1} \subset \mathcal{R} \times \mathcal{R}^{n_i} \times \mathcal{R}^{n_e} \times \mathcal{R}^{2n_c} \times \mathcal{R},$$

where $\lambda^\diamond$ satisfies the Fritz-John conditions (2.14), (2.15).

THEOREM 2.5.

(i) *The formulation (2.13) has a nonempty Lagrange multiplier set if and only if (MPCC) has a nonempty Lagrange multiplier set.*

(ii) *If (MPCC) has a nonempty Lagrange multiplier set and if it satisfies the quadratic growth condition at x^*, then (2.13) has a nonempty Lagrange multipliers set and satisfies the quadratic growth condition at x^*.*

Proof The proof of part **(i)** this result is based on Lemma 2.4, in conjunction with the equations (2.3–2.4) and (2.14–2.15).

The proof of part **(ii)** follows from part **(i)**, the fact that the quadratic growth condition holds at x^*, [7, Theorem 3.113] applied to the L_1 merit function, and the fact that the following inequality holds for x in some neighborhood of x^* and for some $c_{FF} > 0$:

$$\left(\sum_{k=1}^{n_c} F_{k1}(x)F_{k2}(x)\right)^+ \geq \sum_{k=1}^{n_c} \left(F_{k1}(x)F_{k2}(x)\right)^+ - c_{FF}\left(\left|\left|F_1^+(x)\right|\right|_1 + \left|\left|F_2^+(x)\right|\right|_1\right).$$

$\diamond$

Theorems 2.2 and 2.5 give sufficient conditions for (MPCC) and (2.13) to have a nonempty Lagrange multiplier set in spite of the fact that neither satisfy a constraint qualification at any point in the usual sense of nonlinear programming. In Section 3 these conditions will imply that a relaxed version of either (MPCC) or (2.13) will have the same solution as (MPCC) and will satisfy (MFCQ), which makes either approachable by SQP algorithms.

3. The Elastic Mode. An important class of techniques for solving nonlinear programs (1.1) is sequential quadratic programming. The main step in an algorithm of this type is solving the quadratic program

$$\begin{array}{rll} \min_d & \nabla_x \tilde{f}(x)d + d^T \widetilde{W} d, & \\ \text{subject to} & \tilde{g}_i(x) + \nabla_x \tilde{g}_i(x) d \leq 0, & i = 1, 2, \ldots, m \\ & \tilde{h}_j(x) + \nabla_x \tilde{h}_j(x) d = 0, & j = 1, 2, \ldots, r. \end{array} \tag{3.1}$$

The matrix $\widetilde{W}$ can be the Hessian of the Lagrangian (1.1) at (x, λ) [17], where λ is a suitable approximation of a Lagrange multiplier or a positive definite matrix that approximates the Hessian of the Lagrangian on a certain subspace [15, 24, 34]. A trust-region type constraint may be added to (3.1) to enforce good global convergence properties [17]. The solution $\overline{d}$ of (3.1) is then used in conjunction with a merit function and/or line search to determine a new iterate. We give here only a brief description of SQP algorithms, since our interest is solely in showing how the difficulties regarding the potential infeasibility of (3.1) when applied to (MPCC) can be circumvented. For more details about SQP methods see [15, 17, 24, 34].

If a nonlinear program satisfies (MFCQ) at x^* then the quadratic program will be feasible in a neighborhood of x^*. If (MFCQ) does not hold at x^*, however, the possibility exists that (3.1) is infeasible, no matter how close to x^* [30, 36, 40]. This is an issue in the context of this paper because (MPCC) does not satisfy the (MFCQ) at a solution x^*.

In the case of infeasible subproblems some of the SQP algorithms initiate the *elastic mode* [24]. This consists of modifying the nonlinear program (1.1) by relaxing the constraints and adding a penalty term to the objective function.

When the elastic mode is implemented, only the nonlinear constraints are relaxed [24]. To represent this situation in our approach, we assume that $\tilde{g}_i(x)$ for $i = 1, 2, \ldots, l_i$, and $\tilde{h}_j(x)$, for $j = 1, 2, \ldots, l_e$, are linear.

For these constraints, we assume that

[B1] The set $\mathcal{F}_l$ is feasible, where

$$\mathcal{F}_l = \left\{ x \,\middle|\, \tilde{g}_i(x) \leq 0,\ i = 1, 2, \ldots, l_i, \quad \tilde{h}_j(x) = 0, j = 1, 2, \ldots, l_e \right\},$$

[B2] The preceding representation of $\mathcal{F}_l$ is minimal: $\nabla_x \tilde{h}_j(x)$ are linearly independent, $j = 1, 2, \ldots, l_e$, and $\exists d$ such that $\nabla_x \tilde{h}_j(x) d = 0$, $\nabla_x \tilde{g}_i(x) d < 0$.

None of these assumptions induces any loss of generality. Indeed, if $\mathcal{F}_l = \emptyset$, then the original nonlinear program (1.1) is infeasible. Most software for nonlinear programming starts with an analysis of the linear constraints and the infeasibility of the problem, which is the correct outcome when $\mathcal{F}_l = \emptyset$, is immediately detected. Clearly, the interesting situation is when $\mathcal{F}_l$ is feasible, which is our assumption **[B1]**.

If the set $\mathcal{F}_l$ is polyhedral and nonempty, it must have a minimal representation [43, 11]. In addition, this representation can be computed by solving only one linear program [20]. The methods we use in this work are of the sequential quadratic programming type, where all constraints of the nonlinear program are linearized, and the nonlinear constraints are perhaps relaxed. Since the set $\mathcal{F}_l$ is invariant under linearization, any of its representations will produce the same quadratic program subproblems. As long as we do not involve the Lagrange Multipliers of the constraints defining $\mathcal{F}_l$, that are not invariant to a change in representation, assumption **[B2]** does not result in any loss of generality as long as assumption **[B1]** holds.

Define $\tilde{c}_\infty > 0$, x^0, $k = 0$, $\sigma \in (0, \frac{1}{2})$, $\tau \in (0, 1)$, $s > 0$.
QP Find the solution $d = d^k$ of the quadratic program.

$$\begin{array}{lrcl} \text{minimize}_d & \frac{1}{2} d^T d + \nabla_x \tilde{f}(x^k) d & & \\ \text{subject to} & \tilde{h}(x^k) + \nabla_x \tilde{h}(x^k) d & = & 0 \\ & \tilde{g}(x^k) + \nabla_x \tilde{g}(x^k) d & \leq & 0 \end{array}$$

Find the smallest integer $m = m^k$ that satisfies

$$\psi_\infty(x^k + \tau^m s d^k) - \psi_\infty(x^k) \geq \sigma \tau^m s {d^k}^T d^k.$$

Define $x^{k+1} = x^k + \tau^{m_k} s d^k$, and $k = k + 1$.
Go to **QP**.

TABLE 3.1
The model algorithm

Depending on the type of the relaxation, we can have either an L_1 or an L_∞ approach. Our results are related to the situation in which the merit functions $\psi_1(x)$ and $\psi_\infty(x)$ are exact for the nonlinear program (1.1).

Here we consider the case in which the added penalty term is of the L_∞ [3] type:

$$\begin{array}{lrcl} \min_{x,\zeta} & \tilde{f}(x) + c_\infty \zeta & & \\ \text{subject to} & \tilde{g}_i(x) & \leq & 0,\ i = 1, 2, \ldots, l_i, \\ & \tilde{g}_i(x) & \leq & \zeta,\ i = l_i + 1, \ldots, m, \\ & \tilde{h}_j(x) & = & 0,\ j = 1, 2, \ldots, l_e \\ & -\zeta \leq \tilde{h}_j(x) & \leq & \zeta,\ j = l_e + 1, \ldots, r \\ & \zeta & \geq & 0. \end{array} \tag{3.2}$$

All the constraints are now inequality constraints. A quadratic program analogous to (3.1) is constructed for (3.2), which, since **[B1]** and **[B2]** hold, now satisfies (MFCQ) at any feasible point. A feasible point of (3.2), can be immediately obtained by choosing ζ to be sufficiently large.

We make specific claims about one algorithm, presented in Table 3.1. The algorithm, including the line-search rule, is presented in [3, 4]. The algorithm is not necessarily practical, but it serves to show that rates of convergence results can be obtained under very general assumptions. We now define the algorithm for the general nonlinear program (1.1), though we later applied it to (3.2).

For fixed penalty parameter c_∞, the problem (3.2) can be approached by the above SQP algorithms without resulting in an infeasible QP, since the linearization of the problem (3.2) is always feasible if **[B1]** and **[B2]** hold. If for a solution of (3.2) we have that $\zeta = 0$, then the x component of the solution of (3.2) is also a solution of the original, unrelaxed nonlinear program (1.1).

The possibility exists, however, that c_∞ may have to be increased indefinitely before a solution of (1.1) is obtained. In the following theorem we discuss sufficient conditions that ensure that the elastic mode relaxation (3.2) has x^* as a component of the solution for sufficiently large but finite penalty parameter.

THEOREM 3.1. *Assume that assumptions* **[B1]** *and* **[B2]** *hold for (1.1) and that, at a solution x^* of (1.1), we have that*

- *the Lagrange multiplier set of (1.1) is not empty,*
- *the quadratic growth condition (1.16) is satisfied at x^*, and*
- *the data of (1.1) are twice continuously differentiable.*

Then, for sufficiently large but finite value of the penalty parameter c_∞ we have that

1. *$(x^*, 0)$ is a local minimum of (3.2) at which both (MFCQ) and the quadratic growth condition (1.16) are satisfied.*
2. *$(x^*, 0)$ is an isolated stationary point of (3.2).*
3. *If the algorithm in Table 3.1 is initialized sufficiently close to $(x^*, 0)$ with a sufficiently large penalty parameter $\tilde{c}_\infty$, then the sequence x^k of iterates converges R-linearly to x^*.*

Proof We define the fully relaxed nonlinear program

$$\begin{array}{lrcl} \min_{x,\zeta} & \tilde{f}(x) + c_\infty \zeta & & \\ \text{subject to} & \tilde{g}_i(x) & \leq & \zeta, \; i = 1, 2, \ldots, m, \\ & -\zeta \leq \tilde{h}_j(x) & \leq & \zeta, \; j = 1, 2, \ldots, r \\ & \zeta & \geq & 0. \end{array} \tag{3.3}$$

If (x, ζ) is a feasible point of (3.3), it immediately follows from the definition (1.14) of the L_∞ penalty function, $\tilde{P}_\infty(x)$, that $\zeta \geq \tilde{P}_\infty(x)$. From (1.19), under the assumptions stated in this Theorem, we have that there exists $\tilde{c}_\infty > 0$ such that the penalty function $\psi_\infty(x)$ satisfies a quadratic growth condition at x^*. Choose now

$$c_\infty = \tilde{c}_\infty + 1.$$

Using (1.19), we obtain that, in a sufficiently small neighborhood of x^*, we must have that:

$$\tilde{f}(x) + \tilde{c}_\infty \zeta \geq \tilde{f}(x) + \tilde{c}_\infty \tilde{P}_\infty(x) \geq \sigma_1 \left|\left|x - x^*\right|\right|^2 .$$

Whenever $\zeta \leq \frac{1}{\sigma_1}$, we will have that $\sigma_1 \zeta^2 \leq \zeta$. Therefore, in a sufficiently small neighborhood of $(x^*, 0)$, for all (x, ζ) feasible, we will have that:

$$\tilde{f}(x) + c_\infty \zeta = \tilde{f}(x) + \tilde{c}_\infty \zeta + \zeta \geq \sigma_1 \left(\left|\left|x - x^*\right|\right|^2 + \zeta^2 \right).$$

Therefore, for our choice of c_∞ we have that (3.3) satisfies the quadratic growth condition for feasible points (x, ζ). Since any feasible point of (3.2) is feasible for (3.3), if follows that (3.2) also satisfies quadratic growth at $(x^*, 0)$ for every feasible point (x, ζ). Since (3.2) clearly satisfies (MFCQ) everywhere from assumption **[B2]**, this is equivalent to the quadratic growth condition (1.16) holding for all (x, ζ) in a neighborhood of $(x^*, 0)$ [6, 7]. The proof of part 1 of the theorem is complete.

From the conclusion of part 1 we have that, since (MFCQ) and the quadratic growth condition holds for (3.2) at $(x^*, 0)$ this point must be isolated stationary points of the respective nonlinear programs [1]. This concludes the proof of part 2.

Part 3 immediately follows from [1] since (3.2) and satisfies the quadratic growth condition and (MFCQ) at $(x^*, 0)$, Note that $\tilde{c}_\infty$ that enters the definition of ψ_∞ used in the algorithm is not the same as the one in the proof of part 1. It can be shown that now we need $\tilde{c}_\infty \geq c_\infty$, once the latter is chosen. $\diamond$

Discussion

- The same conclusions hold for an L_1 elastic mode, such as the one in SNOPT [24].
- Determining that a solution point is an isolated stationary point is an important issue in nonlinear programming [39, 14, 7]. In practical terms, it means that a nonlinear programming algorithm with global convergence safeguards that does not leave a neighborhood of the solution point x^* will in fact converge to x^*. Example of such algorithms are provided in the references [8, 9, 10, 13, 29, 17].

- A difficulty with the definition of the algorithm is that the successful completion of the algorithm depends on the choice of the parameters c_∞, $\tilde{c}_\infty$, c_1, that need to be sufficiently large but finite. Note that here c_∞ and $\tilde{c}_\infty$ have different purposes: c_∞ is needed to enforce $\zeta = 0$ at a solution of (3.2), whereas $\tilde{c}_\infty$ is the parameter of the merit function $\tilde{\psi}_\infty(x)$ when the preceding algorithm is applied to (3.2). The two parameters are related, but we need at least $\tilde{c}_\infty > c_\infty$ if we follow the proof of linear convergence from [1].
  If (MFCQ) does not hold, the usual update [2, 4], that depends on the Lagrange multipliers, for the value of the penalty parameter cannot be used to adapt c_∞, if the original value is insufficient to result in $\zeta = 0$. An adaptive elastic mode can be implemented where if ζ is not sufficiently small, then the penalty parameters are increased [24]. In Section 4 we show that such an update works under certain conditions.
- Theorem 3.1, part 1 is close in aim to [40, Theorem 8]. Note, however, that the cited result is rather an analysis tool since it involves a nonlinear program whose setup requires the knowledge of the various active sets at the solution. In particular, a nonlinear penalty term appears in the objective function and it involves only the portion of the complementarity constraints coresponding to the complementarity constraints with $k \in \overline{\mathcal{K}}(x^*)$. Such information is not available in an algorithmic framework and is not included in our approach.

We now apply Theorem 3.1 for the case of interest in this work, MPCC. The following corollary is a simple restatement of Theorem 3.1 for (MPCC).

COROLLARY 3.2. *Assume that (MPCC) satisfies the following conditions, at a solution x^*:*

- *The Lagrange multiplier set of (MPCC) not empty. From Corollary 2.3, (SMFCQ) holding for (TNLP) is a sufficient condition for this assumption to hold.*
- *The quadratic growth condition (1.16) is satisfied at x^*.*
- *The data of (MPCC) are twice continuously differentiable.*

Then the conclusions of Theorem 3.1 hold for (MPCC) and for (2.13).

Proof From Theorem 3.1 the conclusion immediately applies for (MPCC). For (2.13) we apply Theorem 2.5 followed by Theorem 3.1 to obtain the conclusion. $\diamond$

Consequently, when started sufficiently close to a solution and with a sufficiently large penalty parameter, the algorithm will converge to that solution of (MPCC) or (2.13) with a sufficiently large but finite c_∞ and $\tilde{c}_\infty$ as soon as (MPCC) satisfies the quadratic growth condition and has a nonempty Lagrange multiplier set at a solution x^*. Since (SMFCQ) is a generic condition for (MPCC) and holds with probability 1 for instances of problems in the MPCC class [40] and the quadratic growth condition is the weakest second-order sufficient condition, this convergence property is expected to hold with probability 1.

3.1. Numerical Experiments. We conducted some numerical experiments on MPCCs from the collection MacMPEC of Sven Leyffer. We used SNOPT [24], an SQP package that implements an adaptive L_1 elastic mode approach.

We considered three types of problem, all of which appear in [36]

1. Stackelberg games [36, Section 12.1], the `gnash` problems.
2. Generalized Nash complementarity points [36, Section 12.2], the `gne` problem, an instance of the problem 12.34 in [36].
3. Optimum packaging problem. with the two-dimensional elliptic operator, discretized on a grid of 8×8, 16×16, and 32×32 elements, which are the

Table 3.2
Results obtained with SNOPT

Problem	Var-Con-CC	Value	Status	Feval	Elastic,Why
gnash14	21-13-1	-0.17904	Optimal	27	Yes, Inf
gnash15	21-13-1	-354.699	Optimal	12	None
gnash16	21-13-1	-241.441	Optimal	7	None
gnash17	21-13-1	-90.7491	Optimal	9	None
gne	16-17-10	0	Optimal	10	Yes, Inf
pack-rig1-8	89-76-1	0.721818	Optimal	15	None
pack-rig1-16	401-326-1	0.742102	Optimal	21	None
pack-rig1-32	1697-1354-1	0.751564	Optimal	19	None

problems `pack-rig` followed by the discretization index in our table.

With the exception of `gne`, all the problems have the complementarity constraints lumped together as one inequality, as in the formulation (2.13).

In the tables showing the results, we indicate the number of variables, constraints, and complementarity constraints ("Var-Con-CC" in the first column), the final value of the objective function, the number of function evaluations and the final status of the run. SNOPT was run through NEOS server [35] at Argonne National Laboratory. We also indicate if and the reason for which the elastic mode was started. In both examples for which the elastic mode was started, this happened because of an infeasible subproblem, a fact that we indicated in the *Elastic, Why* column of Table 3.2 by an "Inf" symbol.

SNOPT solved all the problems presented in a reasonable number of iterations. The fact that the elastic mode was initiated for SNOPT shows that the use of the elastic mode considerably increases the robustness of sequential quadratic programs, since otherwise SNOPT would have failed with an *Infeasible* diagnostic.

The fact that the MPCC problem class is substantially more difficult for nonlinear program solvers, especially interior point solvers was validated by an extensive computational investigation [16]. The elastic mode approach, like the one implemented by SNOPT, is guaranteed to succeed for a finite penalty parameter under the conditions discussed in this paper.

4. A superlinearily convergent algorithm for MPCC. In the following, we present a superlinear convergence result for a special but widely-encountered type of algorithm that uses exact second derivatives. Here we extend the work in [18], in that we relax one assumption that was critical to the convergence proof: that the iterates satisfy either the complementarity constraints or the feasibility conditions for the unrelaxed quadratic program. We will show that in certain distinguished cases, an adaptive elastic mode approach can be used to induce either feasibility or complementarity for all iterates from which superlinear convergence follows from [18].

We assume that the complementary variables are the last components of the unknown vector x. This is not a restrictive assumption: any MPCC can be recast in such a form by using slack variables [18].

Consider the MPCC, that satisfies this assumption for the (2.13) form of (MPCC):

$$
\text{(4.1)} \qquad \text{(MPCCS)} \qquad
\begin{array}{rcccl}
\text{minimize} & f(x) & & & \\
\text{subject to} & g_i(x) & \leq & 0 & i = 1, 2, \ldots, n_i \\
& h_j(x) & = & 0 & j = 1, 2, \ldots, n_j \\
& x_{k1} & \leq & 0 & k = 1, 2, \ldots, n_c \\
& x_{k2} & \leq & 0 & k = 1, 2, \ldots, n_c \\
& \sum_{k=1}^{n_c} x_{1k} x_{2k} & \leq & 0. &
\end{array}
$$

To simplify the subsequent notation, we assume that we are interested in local convergence to a solution x^* that satisfies:

$$
x^*_{k1} = 0, k = 1, 2, \ldots, n_c, \quad x^*_{k2} = 0, k = 1, 2, \ldots, n_d, \quad x^*_{k2} < 0, k = n_{d+1}, \ldots, n_c. \tag{4.2}
$$

We also denote by $x_1 = (x_{11}, x_{21}, \ldots, x_{n_c 1})$ and by $x_2 = (x_{12}, x_{22}, \ldots, x_{n_c 2})$. With this notation, the complementarity constraint becomes $x_1^T x_2 = 0$, or, equivalently over the feasible set of (MPCCS), $x_1^T x_2 \leq 0$.

Note that here we use a more specialized notation convention for the various index sets, same as in Fletcher and al. [18], that is related to the objects defined in (1.22)–(1.26), for which the notation is the same as in [40].

To prove our convergence results, we will invoke stronger conditions than in our preceding sections. One such condition is MPEC-LICQ. We say that (MPCCS) satisfies MPEC-LICQ at x^*, if the associated relaxed nonlinear program satisfies the linear independence constraint qualification (LICQ) at x^*. Specifically, the condition is expressed as:

$$
\text{MPEC-LICQ:} \quad
\begin{array}{l}
\nabla_x g_i(x^*)|_{i \in \mathcal{A}(x^*)}, \ \nabla_x h_j(x^*)|_{j=1,2,\ldots,n_e}, \ e_{k1}|_{k=1,2,\ldots,n_c}, \\
e_{k2}|_{k=1,2,\ldots,n_d}, \ \text{ are linearly independent.}
\end{array}
$$

As opposed to the preceding sections, where a symbol e was used to denote the vector of all ones, in this section we denote by $e_* \in \mathcal{R}^n$ a vector that has zeroes everywhere, except in the $*$ position, where it has a 1. We also denote by $\mathcal{A}(x^*)$ the set of inequality constraints that is active:

$$
\mathcal{A}(x^*) = \{ i \,| i = 1, 2, \ldots, n_i, \ g_i(x^*) = 0 \}.
$$

For an arbitrary point x, we denote by:

$$
\mathcal{A}(x) = \{ i \,| i = 1, 2, \ldots, n_i, \ g_i(x) \geq 0 \}.
$$

The associated (RNLP) is:

$$
\text{(4.3)} \qquad \text{(RNLPS)} \qquad
\begin{array}{rcccl}
\text{minimize} & f(x) & & & \\
\text{subject to} & g_i(x) & \leq & 0 & i = 1, 2, \ldots, n_i \\
& h_j(x) & = & 0 & j = 1, 2, \ldots, n_j \\
& x_{k1} & \leq & 0 & k = 1, 2, \ldots, n_d \\
& x_{k2} & \leq & 0 & k = 1, 2, \ldots, n_d \\
& x_{k1} & = & 0 & k = n_{d+1}, \ldots, n_c.
\end{array}
$$

and, if x^* is a local solution of (MPCCS) and MPEC-LICQ holds at x^*, then x^* is a Karush-Kuhn-Tucker point of (MPCCS) [40] as well as a local solution of (RNLPS). If

MPEC-LICQ holds at x^*, we have that LICQ holds for (RNLPS) at x^*, and (RNLPS) has a unique Lagrange multiplier at the solution: $(\tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta})$, that satisfies

$$\tilde{\nu} \geq 0,\ \tilde{\mu}_{k1} \geq 0,\ \tilde{\mu}_{k2} \geq 0,\ k = 1, 2, \ldots, n_d$$

$$\begin{array}{rcl} -\nabla_x f(x^*) & = & \sum_{i \in \mathcal{A}(x^*)} \tilde{\nu}_i \nabla_x g_i(x^*) + \sum_{i=1}^{n_e} \nabla_x h_j(x^*) \tilde{\pi}_j \\ & + & \sum_{k=1}^{n_d} (\tilde{\mu}_{k1} e_{k1} + \tilde{\mu}_{k2} e_{k2}) + \sum_{k=n_d+1}^{n_c} \tilde{\mu}_{k1} e_{k1}. \end{array}$$

Here we use the natural implicit convention that $\tilde{\nu}_i = 0$ for $i \notin \mathcal{A}(x^*)$, which we will use at other points since this way we can treat the multiplier as a vector that satisfies the complementarity constraints corresponding to the optimality conditions. However, at times we will refer to $\tilde{\nu}$ or similar quantities as to a vector with n_i components, especially in a solution stability context.

When the MPEC-LICQ assumption holds, we can define strong second-order sufficient conditions.

$(MPEC - SOSC)$ $s^T \nabla^2_{xx} \mathcal{L}^* s > 0,\ \forall s \neq 0,\ s \in \mathcal{C}_{RNLPS}$ where $\nabla^2_{xx} \mathcal{L}^*$ is the Hessian of the Lagrangian evaluated at $(x^*, \tilde{\nu}, \tilde{\pi}, \tilde{\mu}, \tilde{\eta})$ and $\mathcal{C}_{RNLPS}$ is the critical cone of (RNLPS) at x^*.

In the rest of this work we invoke the following assumptions,

[A1] f, g, h are twice continuously differentiable.
[A2] (MPCCS) satisfies MPEC-LICQ at the solution x^*.
[A3] (MPCCS) satisfies MPEC-SOSC at the solution x^*.
[A4] We have that $\tilde{\mu}_{k1} > 0$ for $k = 1, 2, \ldots, n_d$.

Note that assumption **[A4]** is equivalent to the following assumption

[A4b] Either $\tilde{\mu}_{k1} > 0$ or $\tilde{\mu}_{k2} > 0$, for $k = 1, 2, \ldots, n_d$.

All one has to do to recover Assumption **[A4]** is to relabel appropriately the variables that appear in complementary pairs. The form of assumption **[A4]** that we consider here will allow us to simplify the notation considerably in our analysis.

At the point in our analysis where we invoke results from [18], we will replace Assumption **[A4]** by a stronger assumption, as well as use an assumption about the solution of the quadratic programs.

[A4a] $\tilde{\nu}_i > 0,\ i \in \mathcal{A}(x^*),\ \tilde{\pi}_j \neq 0,\ j = 1, 2, \ldots, n_e$, and $\tilde{\mu}_{k1} > 0,\ \tilde{\mu}_{k2} > 0$ for $k = 1, 2, \ldots, n_d$.
[A5] When a QP is solved, the QP solver picks a linearly independent basis.

The assumptions **[A1]**–**[A3]** and **[A5]** are identical to the assumptions with the same name used in [18]. Assumption **[A4a]** is identical to Assumption **[A4]** from [18].

Note, however, that at no point in this work we will invoke either Assumption **[A6]** (that the current point satisfies the complementarity constraint $x_1^T x_2 = 0$) or Assumption **[A7]** (that all subproblems are feasible) from the same reference. As we later show, for suitable values of the penalty parameter, the elastic mode approach that we present here will induce one of the situations covered by these two assumptions, without the need to assume it from the outset.

We now explore in more details the relationship between the Lagrange multipliers and the second-order conditions of (RNLPS), (MPCCS), and of an elastic mode relaxation of (MPCCS).

In the following analysis, an important role will be played by the fact that we will assume that the elastic mode implementation enforces the linear constraints exactly. To that end, we assume that:

$$g_i(x),\ i=1,2,\ldots,l_i, \quad h_j(x),\ j=1,2,\ldots,l_e \text{ are linear functions.}$$

This is not a restrictive assumption: it is only a bookkeeping convention, since we allow l_i and l_e to be 0. Also note, following from our assumption **[A2]**, that MPEC-LICQ holds at x^*, that we have that the assumptions **[B1]** and **[B2]** that were used for the elastic mode defined in Section 3, will automatically hold here.

We now present an L_∞ elastic mode approach. An L_1 approach with similar properties can be defined, exactly as used in SNOPT [24]. We use the L_∞ approach because of notational convenience: only one extra variable is needed. In this case, the relaxed MPCC becomes:

$$(MPCC(c)) \qquad \begin{array}{rrcll} \text{minimize} & & f(x)+c_\infty\zeta & & \\ \text{subject to} & & g_i(x) & \leq 0 & i=1,2,\ldots,l_i \\ & & g_i(x) & \leq \zeta & i=l_i+1,\ldots,n_i \\ & & h_j(x) & = 0 & j=1,2,\ldots,l_e \\ & -\zeta\leq & h_j(x) & \leq \zeta & j=l_e+1,\ldots,n_e \\ & & x_{k1} & \leq 0 & k=1,2,\ldots,n_c \\ & & x_{k2} & \leq 0 & k=1,2,\ldots,n_c \\ & & \sum_{k=1}^{n_c} x_{k1}x_{k2} & \leq \zeta & \\ & & \zeta & \geq 0. & \end{array}$$

When MPEC-LICQ holds, we can say more about the relationship between the multipliers of (RNLPS), (MPCCS) and (MPCC(c)), than we did in Section 2. The Lagrange multiplier set of (MPCCS) is not empty at x^*, following [40]. Therefore, there must exist ν_i, $i\in\mathcal{A}(x^*)$; π_j, $j=1,2,\ldots,n_e$; μ_{k1}, for $k=1,2,\ldots,n_c$; and μ_{k2} for $k=1,2,\ldots,n_d$; and $\eta\in\mathcal{R}$, that satisfy the KKT conditions for (MPCCS):

$$\begin{array}{rcl} -\nabla_x f(x^*) & = & \sum_{i\in\mathcal{A}(x^*)}\nu_i\nabla_x g_i(x^*)+\sum_{j=1}^{n_e}\nabla_x h_j(x^*)\pi_j+ \\ & + & \sum_{k=1}^{n_d}(\mu_{k1}e_{k1}+\mu_{k2}e_{k2})+\sum_{k=n_d+1}^{n_c}\mu_{k1}e_{k1}+\eta\sum_{k=n_d+1}^{n_c}x^*_{k2}e_{k1}, \end{array}$$

as well as the inequality constraints $\nu_i\geq 0$, $i\in\mathcal{A}(x^*)$, $\mu_{k1}\geq 0$, $k=1,2,\ldots,n_c$, $\mu_{k2}\geq 0$, $k=1,2,\ldots,n_d$ and $\eta\geq 0$. Here η is the Lagrange multiplier of the complementarity constraint of (MPCCS) $x_1^Tx_2\leq 0$.

One immediate consequence is that (MPCCS) has a Lagrange multiplier that is minimal (in terms of the 1 norm) [18]. We call that multiplier the *fundamental multiplier*. Comparing the algebraic expression of its components in terms of the components of the multiplier of (RNLPS) we obtain:

$$(FMC) \quad \begin{array}{rcll} \nu^* & = & \tilde\nu & \\ \pi^* & = & \tilde\pi & \\ \mu^*_{k1} & = & \tilde\mu_{k1} & \geq 0, \quad k=1,2,\ldots,n_d \\ \mu^*_{k2} & = & \tilde\mu_{k2} & \geq 0, \quad k=1,2,\ldots,n_d \\ \mu^*_{k1} & = & \tilde\mu_{k1}-\eta^*x^*_{k2} & \geq 0, \quad k=n_d+1,\ldots,n_c \\ \eta^* & = & \max\left\{0,\max_{k=n_d+1,\ldots,n_c}\left\{\frac{\tilde\mu_{k1}}{x^*_{k2}}\right\}\right\} & \geq 0. \end{array}$$

Since MPEC-LICQ holds, it is immediate that (MPCCS) has a degeneracy of order 1 and thus its Lagrange multiplier set will have dimension at most 1. Therefore any multiplier of (MPCCS) must satisfy:

$$
\begin{array}{rcll}
\nu & = & \nu^* & \\
\pi & = & \pi^* & \\
\mu_{k1} & = & \tilde{\mu}^*_{k1} \quad \geq 0, & k = 1, 2, \ldots, n_d \\
\mu_{k2} & = & \tilde{\mu}^*_{k2} \quad \geq 0, & k = 1, 2, \ldots, n_d \\
\mu_{k1} & = & \tilde{\mu}^*_{k1} - a x_{k2} \quad \geq 0, & k = n_d + 1, \ldots, n_c \\
\eta & = & \eta^* + a \quad \geq 0, &
\end{array}
\tag{4.4}
$$

where $a \geq 0$.

We now write the KKT conditions for the relaxed problem (MPCC(c)) at $(x^*, 0)$, assuming that $(x^*, 0)$ is a stationary point of (MPCC(c)), an assumption for which we will later determine sufficient conditions for it to hold. We have that, for a Lagrange multiplier of MPCC(c) with components $\hat{\nu}_i$, $i \in \mathcal{A}(x^*)$; $\hat{\pi}_j$, for $j = 1, 2, \ldots, l_e$; $\hat{\pi}^+_j$, associated with the inequality $h_j(x) \leq \zeta$ and $\hat{\pi}^-_j$, associated with the inequality $-h_j(x) \leq \zeta$, $j = l_e + 1, \ldots, n_e$; $\hat{\mu}_{k1}$ for $k = 1, 2, \ldots, n_c$; $\hat{\mu}_{k2}$ for $k = 1, 2, \ldots, n_d$; and $\hat{\eta}$, associated with the complementarity constraint $x_1^T x_2 \leq \zeta$; and $\hat{\theta} \geq 0$, associated with the inequality $\zeta \geq 0$, we can write the KKT conditions:

$$
\begin{array}{rcl}
-\nabla_x f(x^*) & = & \sum_{i \in \mathcal{A}(x^*)} \hat{\nu}_i \nabla_x g_i(x^*) + \sum_{j=1}^{l_e} \nabla_x h_j(x^*) \hat{\pi}_j \\
& + & \sum_{j=l_e+1}^{n_e} \nabla_x h_j(x^*) \left(\hat{\pi}^+_j - \hat{\pi}^-_j \right) + \sum_{k=1}^{n_d} \left(\hat{\mu}_{k1} e_{k1} + \hat{\mu}_{k2} e_{k2} \right) \\
& + & \sum_{k=n_d+1}^{n_c} \hat{\mu}_{k1} e_{k1} + \hat{\eta} \sum_{k=n_d+1}^{n_c} \left(x^*_{k2} e_{k1} \right) \\
c_\infty & = & \sum_{i=l_i+1,\ldots,n_i,\; i \in \mathcal{A}(x^*)} \hat{\nu}_i + \sum_{j=l_e+1}^{n_e} \left(\hat{\pi}^+_j + \hat{\pi}^-_j \right) + \hat{\eta} + \hat{\theta}.
\end{array}
$$

If MPEC-LICQ holds, and $(x^*, 0)$ is a stationary point of (MPCC(c)), then, using (4.4), we obtained that any Lagrange multiplier of (MPCC(c)) must satisfy, in terms of the components of the fundamental multiplier, the following relations:

$$
\begin{array}{rcllll}
\hat{\nu}_i & = & \nu^*_i & \geq 0 & i \in \mathcal{A}(x^*) \\
\hat{\pi}_j & = & \pi^*_j & & j = 1, 2, \ldots, l_e, \\
\hat{\pi}^+_j & = & \max\{\pi^*_j, 0\} + f_j & \geq 0 & j = l_e + 1, \ldots, n_e \\
\hat{\pi}^-_j & = & \max\{-\pi^*_j, 0\} + f_j & \geq 0 & j = l_e + 1, \ldots, n_e \\
\hat{\mu}_{k1} & = & \mu^*_{k1} & \geq 0 & k = 1, 2, \ldots, n_d \\
\hat{\mu}_{k2} & = & \mu^*_{k2} & \geq 0 & k = 1, 2, \ldots, n_d \\
\hat{\mu}_{k1} & = & \mu^*_{k1} - a x_{k2} & \geq 0 & k = n_d + 1, \ldots, n_c \\
\hat{\eta} & = & \eta^* + a & \geq 0 & \\
\hat{\theta} & = & c_\infty - \sum_{i=l_{i+1}, i \in \mathcal{A}(x^*)}^{n_i} \hat{\nu}^*_i & & \\
& - & \sum_{j=l_e+1}^{n_e} \left(\hat{\pi}^+_j + \hat{\pi}^-_j \right) - \hat{\eta} & \geq 0, &
\end{array}
\tag{4.5}
$$

where $a \geq 0$, and $f_j \geq 0$, for $j = l_e + 1, \ldots, n_e$. Here $\hat{\theta}$ is the Lagrange multiplier of the constraint $\zeta \geq 0$ and the requirement $\hat{\theta} \geq 0$ results in a condition on the value that c_∞ must take for $(x^*, 0)$ to be a stationary point of (MPCC(c)). We do not add the complementarity constraints that appear when applying the first order optimality condition, since these are automatically satisfied by our choice of active set and multiplier components.

Rewriting $\hat{\theta}$ in terms of the fundamental multiplier (FMC), we obtain that:

$$
\hat{\theta} = c_\infty - \sum_{i=l_{i+1}, i \in \mathcal{A}(x^*)}^{n_i} \nu^*_i - \sum_{j=l_e+1}^{n_e} \left(|\pi^*_j| + 2 f_j \right) - \eta^* - a \geq 0. \tag{4.6}
$$

We denote by ν_0 the following quantity, that is also defined in terms of the components of the fundamental multiplier (FMC):

$$\nu_0 = \sum_{i=l_{i+1}, i\in\mathcal{A}(x^*)}^{n_i} \nu_i^* + \sum_{j=l_e+1}^{n_e} |\pi_j^*| + \eta^*. \tag{4.7}$$

We have the following Lemma.

LEMMA 4.1. *If $c_\infty \geq \nu_0$, where ν_0 is given by (4.7), then*

1. *The point $(x^*, 0)$ is a stationary point of (MPCC(c)).*
2. *The set of Lagrange multipliers of (MPCC(c)) is not empty and is defined by (4.5), where $a \geq 0$ and $f_j \geq 0$, $j = l_e+1, \ldots, n_e$, also satisfy*

$$c_\infty - \nu_0 \geq \sum_{j=l_e+1}^{n_e} (2f_j) + a. \tag{4.8}$$

Proof Since the values of the fundamental multiplier are fixed, and since $a \geq 0$ and $f_j \geq 0$, $j = l_e+1, \ldots, n_e$, the necessary and sufficient condition for $(x^*, 0)$ to be a stationary point of (MPCC(c)) is that $\hat{\theta} \geq 0$. From equations (4.6) and (4.7) this is equivalent to

$$c_\infty \geq \nu_0,$$

which proves the first part of the claim.

By inspecting (4.5) we see that the set of Lagrange multipliers of (MPCC(c)) is defined by $f_j \geq 0$, $j = l_e+1, \ldots, n_e$ and $a \geq 0$ that satisfy the inequality $\hat{\theta} \geq 0$, that is,

$$c_\infty - \nu_0 \geq \sum_{j=l_e+1}^{n_e} (2f_j) + a.$$

The proof is complete. $\diamond$

An immediate consequence is that, if $c_\infty = \nu_0$, then the program (MPCC(c)) has a unique multiplier! This is formally stated in the next result.

4.1. Second-order conditions. LEMMA 4.2. *Assume that (MPCCS) satisfies MPEC-LICQ and MPEC-SOSC. Assume that c_∞ satisfies $c_\infty \geq \nu_0$. Then (MPCC(c)) satisfies (MFCQ) and (RSOSC) at $(x^*, 0)$. In addition, if $c_\infty = \nu_0$, then the Lagrange multiplier set of (MPCC(c)) at $(x^*, 0)$ has a unique element.*

Proof Consider one multiplier of (MPCC(c)), whose components are: $\hat{\nu}_i \geq 0$ for $i \in \mathcal{A}(x^*)$; $\hat{\pi}_j$ for $j = 1, 2, \ldots, l_e$; $\hat{\pi}_j^+ \geq 0$ and $\hat{\pi}_j^- \geq 0$ for $j = l_e+1, \ldots, n_e$; $\hat{\mu}_{k1} \geq 0$ and $\hat{\mu}_{k2} \geq 0$ for $k = 1, 2, \ldots, n_d$; $\hat{\mu}_{k1} \geq 0$ for $k = 1, 2, \ldots, n_d$; $\hat{\eta} \geq 0$ and $\hat{\theta} \geq 0$.

The Hessian of the Lagrangian at x^* corresponding to this Lagrange Multiplier, accounting for only the nonlinear terms, and using (FMC) and (4.5), is the following:

$$\begin{array}{rcl}
\nabla^2_{xx}\mathcal{L}_{MPCC(c)}(x^*, 0) & = & \nabla^2_{xx} f(x^*) + \sum_{j=l_e+1}^{n_e} \left(\hat{\pi}_j^+ \nabla^2_{xx} h_j(x^*) - \hat{\pi}_j^- \nabla^2_{xx} h_j(x^*)\right) \\
& + & \sum_{i=l_i+1, i\in\mathcal{A}(x^*)}^{n_i} \hat{\nu}_i \nabla^2_{xx} g(x^*) + \hat{\eta} \sum_{k=1}^{n_c} \left(e_{k_1}^T e_{k2} + e_{k2}^T e_{k1}\right) \\
& = & \nabla^2_{xx} f(x^*) + \sum_{j=l_e+1}^{n_e} \pi_j^* \nabla^2_{xx} h_j(x^*) \\
& + & \sum_{i=l_i+1, i\in\mathcal{A}(x^*)}^{n_i} \nu_i^* \nabla^2_{xx} g(x^*) + \hat{\eta} \sum_{k=1}^{n_c} \left(e_{k_1}^T e_{k2} + e_{k2}^T e_{k1}\right) \\
& = & \nabla^2_{xx}\mathcal{L}_{RNLPS}(x^*) + \hat{\eta} \sum_{k=1}^{n_c} \left(e_{k_1}^T e_{k2} + e_{k2}^T e_{k1}\right).
\end{array}$$

The Hessian of the Lagrangian of (MPCC(c)) should also be computed with respect to ζ, but, since the contribution of ζ is linear both in the constraints and in the objective function of (MPCC(c)), it follows that the Hessian of the Lagrangian is:

$$\nabla^2_{(x,\zeta)(x,\zeta)}\mathcal{L}_{MPCC(c)}(x^*,0) = \begin{pmatrix} \nabla^2_{xx}\mathcal{L}_{MPCC(c)}(x^*,0) & 0_{n\times 1} \\ 0_{1\times n} & 0 \end{pmatrix}.$$

Note that, from Lemma 2.1, we have that $\mathcal{C}_{MPCCS} = \mathcal{C}_{RNLPS}$ at x^*, where by $\mathcal{C}$ we denote the critical cone of the respective nonlinear program. When $c_\infty \geq \nu_0$, since x^* is a stationary point for (MPCC) and $(x^*,0)$ is a stationary point for (MPCC(c)) we will have that $\mathcal{C}_{MPCC(c)} = \mathcal{C}_{MPCCS} \bigoplus \{0\}$. Let now $(u,0) \in \mathcal{C}_{MPCC(c)}$ and thus $u \in \mathcal{C}_{RNLPS}$. We have that:

$$\begin{aligned} (u,0)^T\nabla^2_{(x,\zeta)(x,\zeta)}\mathcal{L}_{MPCC(c)}(u,0) &= u^T\mathcal{L}_{MPCC(c)}u = u^T\mathcal{L}_{RNLPS}u \\ &+ 2\hat{\eta}\textstyle\sum_{k=1}^{n_c}(e_{k1}^Tu)(e_{k2}^Tu) \geq u^T\mathcal{L}_{RNLPS}u > 0, \end{aligned}$$

where the last two inequalities follow, respectively, from the fact that, on the critical cone of (RNLPS) we have $(e_{k1}^Tu)(e_{k2}^Tu) \geq 0$, for $k = 1, 2, \ldots, n_d$, and $(e_{k1}^Tu)(e_{k2}^Tu) = 0$, for $k = n_d+1, \ldots, n_c$, and $\hat{\eta} \geq 0$, and, respectively, from the MPEC-SOSC assumption. Since we obtained, for any multiplier of MPCC(c), that:

$$\forall \hat{u} \in \mathcal{C}_{MPCC(c)},\ \hat{u} \neq 0 \quad \Rightarrow \quad \hat{u}^T\nabla^2_{(x,\zeta)(x,\zeta)}\mathcal{L}_{MPCC(c)}(x^*,0)\hat{u} > 0,$$

this means that Robinson's condition (RSOSC) holds for (MPCC(c)) at $(x^*,0)$. Since (MFCQ) clearly holds at $(x^*,0)$, the conclusion of the lemma follows. The uniqueness of the multiplier for the case $c_\infty = \nu_0$ follows from (4.8). $\diamond$

4.2. The algorithm. We now consider the effect of applying SQP to MPCC in either the original or the relaxed form. As matrix W of the QP we use the Hessian of the Lagrangian, with the Lagrange multipliers computed at the previous step.

We define the following matrix $\widehat{W}$ at the point x:

$$\widehat{W} = \begin{cases} \nabla^2_{xx}f(x) + \sum_{i=l_i+1}^{n_i}\hat{\nu}_i\nabla^2_{xx}g_i(x) \\ \quad + \sum_{j=l_e+1}^{n_e}\left(\hat{\pi}_j^+ - \hat{\pi}_j^-\right)\nabla^2_{xx}h_j(x) \\ \quad + \hat{\eta}\sum_{k=1}^{n_c}\left(e_{k1}^Te_{k2} + e_{k2}^Te_{k1}\right). \\ \text{or, alternatively,} \\ \nabla^2_{xx}f(x) + \sum_{i=l_i+1}^{n_i}\nu_i\nabla^2_{xx}g_i(x) \\ \quad + \sum_{j=l_e+1}^{n_e}\pi_j\nabla^2_{xx}h_j(x) \\ \quad + \eta\sum_{k=1}^{n_c}\left(e_{k1}^Te_{k2} + e_{k2}^Te_{k1}\right), \end{cases}$$

where the first branch is followed in the quadratic program associated with the relaxation (QPMC(c)) was solved at the preceding iteration, and the second branch is followed if the quadratic program associated with the unrelaxed problem ((QPX), to be defined later) was solved at the preceding iteration.

Consider now the following quadratic program associated to (MPCC(c)) at the point (x,ζ).

$$(QPMC(c))\left\{\begin{array}{lrcll} \min_{d,d_\zeta} & \frac{1}{2}d^T\widehat{W}d+\nabla_x f(x)d+c_\infty(\zeta+d_\zeta) & & & \\ & g_i(x)+\nabla_x g_i(x)d & \leq & 0 & i=1,2,\ldots,l_i \\ & g_i(x)+\nabla_x g_i(x)d & \leq & \zeta+d_\zeta & i=l_i+1,\ldots,n_i \\ & h_j(x)+\nabla_x h_j(x)d & = & 0 & j=1,2,\ldots,l_e \\ & h_j(x)+\nabla_x h_j(x_k)d & \leq & \zeta+d_\zeta & j=l_{e+1},\ldots,n_e \\ \text{sbj. to} & -h_j(x)-\nabla_x h_j(x_k)d & \leq & \zeta+d_\zeta & j=l_{e+1},\ldots,n_e \\ & x_{k1}+d_{k1} & \leq & 0 & k=1,2,\ldots,n_c \\ & x_{k2}+d_{k2} & \leq & 0 & k=1,2,\ldots,n_c \\ & x_1^Td_2+x_2^Td_1+x_1^Tx_2 & \leq & \zeta+d_\zeta & \\ & \zeta+d_\zeta & \geq & 0. & \end{array}\right.$$

To obtain results for the application of the algorithm to (MPCCS), we state and prove the following stability result for general nonlinear programming.

Lemma 4.3. *Consider the nonlinear program:*

$$\begin{array}{rrcl} min_x & \tilde{f}(x) & & \\ subject\ to & \tilde{h}_j(x) & = & 0,\ j=1,2,\ldots,r \\ & \tilde{g}_i(x) & \leq & 0,\ i=1,2,\ldots,m. \end{array}$$

Assume that it satisfies (MFCQ) and (RSOSC) at a solution x^*. We denote the compact Lagrange multiplier set of this nonlinear program by $\Lambda(x^*)$.

Consider the Quadratic Program:

$$\begin{array}{rrcl} min_d & \nabla_x\tilde{f}(x)d+\frac{1}{2}d^T\tilde{W}d & & \\ subject\ to & \tilde{g}_i(x)+\nabla_x\tilde{g}_i(x)d & \leq & 0,\ i=1,2,\ldots,m \\ & \tilde{h}_j(x)+\nabla_x\tilde{h}_j(x)d & = & 0,\ j=1,2,\ldots,r, \end{array}$$

where $\tilde{W}=\nabla^2_{xx}\tilde{f}+\sum_{j=1}^{n_e}\pi_j\nabla^2_{xx}\tilde{h}_j(x)+\sum_{i=1}^{n_i}\nu_i\nabla^2_{xx}\tilde{g}_i(x)$. Let d be a solution of this quadratic program, and ν^+ and π^+ its Lagrange multipliers.

Then, there exist $\epsilon>0$, $c_1>0$, $c_2>0$ such that $\operatorname{dist}((\nu,\pi);\Lambda(x^*))\leq\epsilon$ *and* $\operatorname{dist}(x;x^*)\leq\epsilon$ *imply that*

- **(i)** $\frac{1}{c_1}||x-x^*||\leq||d||\leq c_2||x-x^*||$,
- **(ii)** $\operatorname{dist}((\nu^+,\pi^+);\Lambda(x^*))\leq c_2||x-x^*||$.

Here $\operatorname{dist}(\cdot,\cdot)$ *denotes the distance between two sets.*

Proof The rightmost inequalities in both part **(i)** and **(ii)** are a consequence of [45, Theorem A1]. In that reference, the inequality-only case is treated, but the conclusion can be immediately extended for the case where there are also equality constraints, that are linearly independent and that, together with the inequality constraints, satisfy (MFCQ). Since (MFCQ) holds, we have that:

$$\begin{array}{rcl} \tilde{g}(x+d) & \leq & O(||d||^2)\leq O(||x-x^*||^2) \\ \tilde{h}(x+d) & \leq & O(||d||^2)\leq O(||x-x^*||^2). \end{array}$$

This implies that the size of the infeasibility, $\mathcal{I}(x)$, where

$$\mathcal{I}(x)=\max\left\{\max_{i=1,2,\ldots,n_i}\{\tilde{g}_i^+(x)\},\max_{j=1,2,\ldots,n_e}\{|\tilde{h}_j(x)|\}\right\}$$

satisfies $\mathcal{I}(x) = O(||d||)$. Using the result [45, Theorem A1], as well as $\nu^+ \geq 0$ and the optimality conditions for the QP, we obtain that, since $Wd = -\nabla_x \tilde{f}^T(x) - \nabla_x \tilde{g}^T(x)\nu^+ - \nabla_x \tilde{h}^T(x)\pi^+$, we have that:

$$\begin{aligned}&\Omega\left(||x-x^*|| + \operatorname{dist}\left((\nu^+,\pi^+);\Lambda(x^*)\right)\right) = \\ &\max\left\{\left|\left|\nabla_x \tilde{f}(x) + \nabla_x \tilde{g}(x)\nu^+ + \nabla_x \tilde{h}(x)\pi^+\right|\right|, \mathcal{I}(x)\right\} = O(||d||).\end{aligned}$$

In turn, this implies that there exists a $c_1 > 0$ such that $||x - x^*|| \leq c_1\, ||d||$, which completes the proof. $\diamond$

In the following, we will characterize the behavior of the solution of (QPMC(c)). Towards that end, we considered the following multipliers. The multiplier

$$\widehat{\Pi} = \left(\widehat{\nu}, \widehat{\pi}, \widehat{\pi}^+, \widehat{\pi}^-, \widehat{\mu}_1, \widehat{\mu}_2, \widehat{\nu}, \widehat{\theta}\right)$$

represents an approximation to the multiplier of the quadratic program (QPMC(c)). The multiplier

$$\widehat{\Pi}^+ = \left(\widehat{\nu}^+, \widehat{\pi}^+, \widehat{\pi}^{+^+}, \widehat{\pi}^{-^+}, \widehat{\mu}_1^+, \widehat{\mu}_2^+, \widehat{\nu}^+, \widehat{\theta}^+\right)$$

represents the multiplier of the quadratic program (QPMC(c)) computed at (x,ζ), with the approximation to the Hessian matrix, $\widehat{W}$, computed with $\widehat{\Pi}$. Finally, the multiplier

$$\widehat{\Pi}^* = \left(\widehat{\nu}^*, \widehat{\pi}^*, \widehat{\pi}^{+^*}, \widehat{\pi}^{-^*}, \widehat{\mu}_1^*, \widehat{\mu}_2^*, \widehat{\nu}^*, \widehat{\theta}^*\right)$$

is a Lagrange multiplier of (MPCCS) at $(x^*, 0)$.

Lemma 4.4. *Consider the quadratic program (QPMC(c)), whose solution we denote by (d, d_ζ). Assume that $c_\infty - \nu_0 = \gamma_0 > 0$. There exists $\epsilon > 0$ such that, if*

$$\operatorname{dist}\left(\widehat{\Pi}, \Lambda_{MPCC(c)}\right) < \epsilon \quad \textit{and} \quad \operatorname{dist}\left((x,\zeta),\ (x^*,0)\right) < \epsilon,$$

then the following hold.

(i) If $\zeta + d_\zeta > 0$, then $x + d$ satisfies $x_{k1} + d_{k1} = 0$, for $k = 1, 2, \ldots, n_c$.

(ii) If $x_{k1} = 0$, for $k = 1, 2, \ldots, n_c$, then $\zeta + d_\zeta = 0$, and $x + d$ also satisfies $x_{k1} + d_{k1} = 0$, for $k = 1, 2, \ldots, n_c$.

Note A point that satisfies $x_{k1} = 0$, for $k = 1, 2, \ldots, n_c$, also satisfies the complementarity constraints $x_1^T x_2 = 0$.

Proof of part (i) Assume that $\zeta + d_\zeta > 0$. Using Lemma 4.2 and Lemma 4.3, it follows from the expression of the Lagrange multipliers of (MPCC(c)) at $(x^*, 0)$, (4.5), that we must have that

$$\widehat{\mu}_{k1}^+ > 0 \quad k = 1, 2, \ldots, n_d. \tag{4.9}$$

The last result follows from Assumption [**A4**].

As a result we obtain from the optimality conditions of (QPMC(c)) that

$$x_{k1} + d_{k1} = 0, \quad k = 1, 2, \ldots, n_d. \tag{4.10}$$

In addition, since $\zeta + d_\zeta > 0$, not both bound constraints on $h_j(x) + \nabla_x h_j(x) d$ in (QPMC(c)) can be active at the same time. Using the optimality conditions of

(QPMC(c)), this implies that $\widehat{\pi}_j^{+^+}\widehat{\pi}_j^{-^+} = 0$ for $j = n_{l_e+1}, \ldots, n_e$. Assume that $\widehat{\pi}_j^{-^+} = 0$ for some $n_{l_e+1} \leq j \leq n_e$.

Let now $\widehat{\Pi}^*$ be a Lagrange multiplier of (MPCC(c)) that satisfies $\left|\left|\widehat{\Pi}^* - \widehat{\Pi}\right|\right| \leq c_2\left(||x - x^*|| + |\zeta - \zeta^*|\right)$, where $\zeta^* = 0$. Such a multiplier must exist by Lemma 4.2 and Lemma 4.3. It then follows that $0 \leq \widehat{\pi}_j^{-^*} \leq c_2\left(||x - x^*|| + |\zeta|\right)$. Define now $f_j^* = \widehat{\pi}_j^{-^*} - \max\{-\pi_j^*, 0\}$. From the preceding two equations we must have that $f_j^* \leq c_2\left(||x - x^*|| + \zeta\right)$, whereas from (4.5) we must have that $f_j^* \geq 0$.

Applying this argument for any $n_{l_e+1} \leq j \leq n_e$, we obtain that

$$\sum_{j=l_e+1}^{n_e} f_j^* \leq c_2 n_e\left(||x - x^*|| + \zeta\right).$$

Finally, $\zeta + d_\zeta > 0$ implies, from the optimality conditions of (QPMC(c)), that $\widehat{\theta}^+ = 0$, and thus $0 \leq \widehat{\theta}^* \leq c_2\left(||x - x^*|| + \zeta\right)$.

Define now $a^* = \widehat{\eta}^* - \eta^*$. Using (4.5) and the fact that $c_\infty - \nu_0 = \gamma_0 > 0$, we obtain that

$$a^* = \gamma_0 - \sum_{j=l_e+1}^{n_e} 2f_j^* - \widehat{\theta}^* \geq \gamma_0 - (2n_e + 1)c_2\left(||x - x^*|| + \zeta\right).$$

It therefore follows that, since $\gamma_0 > 0$, if we choose $\epsilon > 0$ sufficiently small we have that $a^* \geq \frac{\gamma_0}{2}$. Using again (4.5), Lemmas 4.2 and 4.3, as well as our definition of $\widehat{\Pi}^*$, we obtain that

$$\widehat{\mu}_{k1}^+ > 0, \quad k = n_d + 1, \ldots, n_c.$$

From the optimality conditions of (QPMC(c)) and (4.9), this implies that

$$x_{k1} + d_{k1} = 0, \quad k = 1, 2, \ldots, n_c.$$

Thus the proof of the first part of the result is complete.

Proof of part (ii) Assume now that x satisfies $x_{k1} = 0$, $k = 1, 2, \ldots, n_c$. We define

$$W = \widehat{W} - \widehat{\eta}\sum_{k=1}^{n_c}(e_{k1}^T e_{k2} + e_{k2}^T e_{k1}).$$

Consider the quadratic program

$$(QPR)\left\{\begin{array}{llcll} \min_d & \frac{1}{2}d^T W d + \nabla_x f(x) d & & & \\ & g_i(x) + \nabla_x g_i(x) d & \leq & 0 & i = 1, 2, \ldots, n_i \\ \text{sbj. to} & h_j(x) + \nabla_x h_j(x) d & = & 0 & j = 1, 2, \ldots, n_e \\ & d_{k1} & = & 0 & k = 1, 2, \ldots, n_c \\ & x_{k2} + d_{k2} & \leq & 0 & k = 1, 2, \ldots, n_c \end{array}\right.$$

Using Assumption **[A2]**, it is immediate that (QPR) satisfies LICQ in a neighborhood of $x = x^*$ and $d = 0$. Using Lemma 4.2 and 4.3 it follows that as $x \to x^*$, (QPR) has a unique solution in the neighborhood of $d = 0$, and the unique corresponding Lagrange multiplier $\tilde{\Pi}_R$ of (QPR) approaches the unique Lagrange multiplier $\tilde{\Pi}$ of (RNLPS) at $x = x^*$.

Also, since we have that $d^T\widehat{W}d = d^TWd$ for any d in the feasible set of (QPR), it follows that the quadratic program that is obtained by substituting W by $\widehat{W}$ has the same solution and the same multiplier.

It then follows that d is also a stationary point of the following quadratic program

$$(QPX)\begin{cases} \min_d & \frac{1}{2}d^T\widehat{W}d + \nabla_x f(x)d \\ \text{sbj. to} & \begin{array}{rcll} g_i(x)+\nabla_x g_i(x)d & \leq & 0 & i=1,2,\ldots,n_i \\ h_j(x)+\nabla_x h_j(x)d & = & 0 & j=1,2,\ldots,n_e \\ x_{k1}+d_{k1} & \leq & 0 & k=1,2,\ldots,n_c \\ x_{k2}+d_{k2} & \leq & 0 & k=1,2,\ldots,n_c \\ x_1^Td_2+x_2^Td_1+x_1^Tx_2 & \leq & 0. & \end{array} \end{cases}$$

with a multiplier Π that is constructed from the multiplier $\tilde{\Pi}_R$ by the same procedure that we used to construct the fundamental multiplier (FMC) from $\tilde{\Pi}$. It therefore follows that $\Pi = (\nu, \pi, \mu_{k1}, \mu_{k2}, \eta)$ approaches the fundamental multiplier defined in (FMC) as x approaches x^*.

It then follows that we can construct from Π a Lagrange multiplier $\widehat{\Pi}$ for (QPMC(c)) by the same relation that we used in (4.5) to construct all Lagrange multipliers of (MPCC(c)) starting from the fundamental multiplier of (MPCCS), where in the same procedure we take $a = 0$.

The only relation in (4.5) that is not immediately true is $\widehat{\theta} \geq 0$. However, since we start from a multiplier that is close to the fundamental multiplier, since we take $a = 0$, and since we assume that $c_\infty - \nu_0 = \gamma_0 > 0$, it is clear that $\widehat{\theta}$ approaches γ_0 and is therefore positive when $\epsilon > 0$ is sufficiently small.

It then follows that we can choose $d_\zeta = -\zeta$ and then (d, d_ζ) is a stationary point of (QPMC(c)). However, from Lemma 4.2, it follows from [39] that the solution of (QPMC(c)) in a neighborhood of $(0,0)$ is local unique. Therefore (d, d_ζ) is a solution of (QPMC(c)). Since, following the definition of (QPR), for that solution we will have that $x_{k1} + d_{k1} = 0$, $k = 1, 2, \ldots, n_c$ as well as $\zeta + d_\zeta = 0$, the conclusion follows. $\diamond$

4.3. A superlinearily convergent algorithm. We now state our algorithm for general nonlinear programming, and then show that, under certain conditions, it will converge superlinearily when applied to (MPCCS). The problem to be solved is the following:

$$\text{(4.11)} \qquad \begin{array}{rlcll} \min_x & \tilde{f}(x) & & & \\ \text{sbj. to} & \tilde{g}_i(x) & \leq & 0, & i=1,2,\ldots,m \\ & \tilde{h}_j(x) & = & 0, & j=1,2,\ldots,r. \end{array}$$

We assume that $\tilde{g}_i(x)$, $i = 1, 2, \ldots, l_i$ and $\tilde{h}_j(x)$, $j = 1, 2, \ldots, l_e$, are linear constraints that do not get relaxed in an elastic mode approach.

In the algorithm that we describe we use one of the two following quadratic programs:

$$(QP) \quad \begin{array}{rlcll} \min_d & \nabla_x\tilde{f}(x)d + \frac{1}{2}d^T\tilde{W}d & & & \\ \text{sbj. to} & \tilde{g}_i(x)d + \nabla_x\tilde{g}_i(x)^Td & \leq & 0, & i=1,2,\ldots,m \\ & \tilde{h}_j(x)d + \nabla_x\tilde{h}_j(x)^Td & = & 0, & j=1,2,\ldots,r \end{array}$$

$x^0 = x$, $c_\infty = c_0$, $k = 0$.
NLP: Solve (QP).
If $\sum_{i=l_i+1}^{m} \nu_i^q + \sum_{j=l_e+1}^{r} |\pi_j^q| \leq c_\mu$ and (QP) is feasible
    $x^{q+1} = x^q + d^q$, $q = q + 1$, return to **NLP**.
Else
    NLPC: solve (QPC).
    $x^{q+1} = x^q + d^q$, $\zeta^{q+1} = \zeta^q + \delta_\zeta^q$, $q = q + 1$.
    If $\sqrt{||d^q|| + \left|\left|\delta_\zeta^q\right|\right|} \leq \zeta^n$,
        $c_\infty = c_\infty c_\gamma$, $q = q + 1$ return to **NLPC**.
    End If
End If

TABLE 4.1
The elastic mode algorithm

$$
(QPC) \quad \begin{array}{rrcll}
\min_{d,d_\zeta} & \nabla_x \tilde{f}(x) d + \frac{1}{2} d^T \tilde{W} d & + & c_\infty(\zeta + d_\zeta) & \\
\text{sbj. to} & \tilde{g}_i(x) d + \nabla_x \tilde{g}_i(x) d & \leq & 0, & i = 1, 2, \ldots, l_i \\
 & \tilde{g}_i(x) d + \nabla_x \tilde{g}_i(x) d & \leq & \zeta + d_\zeta, & i = l_i + 1, \ldots, m \\
 & \tilde{h}_j(x)^T d + \nabla_x \tilde{h}_j(x) d & = & 0, & j = 1, 2, \ldots, l_e \\
-\zeta - d_\zeta \leq & \tilde{h}_j(x)^T d + \nabla_x \tilde{h}_j(x) d & \leq & \zeta + d_\zeta, & j = l_e + 1, \ldots, r \\
 & \zeta + d_\zeta & \geq & 0. &
\end{array}
$$

When (QP) is solved, we obtain a direction d and Lagrange multipliers ν_i, $i = 1, 2, \ldots, n_i$ and π_j, for $j = 1, 2, \ldots, n_e$. When (QPC) is solved, we obtain Lagrange multipliers ν_i, $i = 1, 2, \ldots, n_i$; π_j, $j = 1, 2, \ldots, l_i$; π_j^+ and π_j^-, $j = l_e + 1, \ldots, n_j$; and θ. We define the matrix $\tilde{W}$ to be used to in the next quadratic program, which is either (QP) or (QPC) as follows.

$$
\tilde{W} = \begin{cases}
\nabla_{xx}^2 \tilde{f}(x) + \sum_{i=l_i+1}^{r} \nu_i \nabla_{xx}^2 \tilde{g}_i(x) + \sum_{j=l_e+1}^{m} \pi_j \nabla_{xx}^2 \tilde{h}_j(x), \\
\qquad \text{if (QP) was last solved} \\
\nabla_{xx}^2 \tilde{f}(x) + \sum_{i=l_i+1}^{r} \nu_i \nabla_{xx}^2 \tilde{g}_i(x) + \sum_{j=l_e+1}^{m} \left(\pi_j^+ - \pi_j^-\right) \nabla_{xx}^2 \tilde{h}_j(x), \\
\qquad \text{if (QPC) was last solved.}
\end{cases}
$$

We now define our algorithm in Table 4.1. The algorithm depends upon the parameters c_μ, $c_\gamma > 1$, c_0.

The (QP) subproblem of this algorithm is the same subproblem as in the algorithm by [19], minus the trust-region constraint that is imposed for globalization. The (QPC) subproblem is the natural extension of (QP) when using the elastic mode. The elastic mode strategy used here is identical to the one used in SNOPT [24], except that we use here the L_∞ function.

Note that once the subproblem (QPC) is solved the algorithm never solves the problem (QP) again. So we either solve (QP) till convergence or (QPC) till convergence.

We now analyze the effect of applying this algorithm to (MPCC). For this case, the (QP) subproblem becomes (QPX), whereas the (QPC) subproblem becomes (QPMC(c)).

THEOREM 4.5. *Assume that Assumptions* **[A1]**–**[A3]**, **[A4a]** *and* **[A5]** *hold near a solution* x^* *of (MPCCS). Assume that the point* x^{q_0} *is sufficiently close to* x^* *and either*

- *(i) The elastic mode is never invoked, and the algorithm uses at* q_0, *for the purpose of constructing the matrix* $\tilde{W}$, *an estimate of the Lagrange multiplier that is sufficiently close to a multiplier of (MPCCS).*
- *(ii) The elastic mode is invoked at* q_0 *with* $c_\infty \geq \nu_0$ *and at all subsequent iterates, and the algorithm uses at* q_0, *for the purpose of constructing the matrix* $\tilde{W}$, *an estimate of the Lagrange multiplier that is sufficiently close to a multiplier of (MPCC(c)), and a* ζ^{q_0} *that is sufficiently close to* 0.

Then x^q *converges to* x^* *superlinearly in case* **(i)** *and* (x^q, ζ^q) *converges to* $(x^*, 0)$ *superlinearly in case* **(ii)**.

Proof For case **(i)** to hold, we must have that the test involving the 1 norm of the nonlinear multipliers is always satisfied. Therefore, for all iteration indices q, we have that the following inequality holds:

$$\sum_{i=l_i+1}^{n_i} \nu_i^q + \sum_{j=l_e+1}^{n_e} |\pi_j^q| + \eta^q \leq c_\mu,$$

where the left hand side is composed of multipliers of (QPX). Since we are sufficiently close to x^*, this must imply that $c_\mu \geq \nu_0$. In the latter case, we will have that a solution of (QPX) can be completed to a solution of (QPMC(c)) with $\zeta = 0$ and $\zeta + \delta_\zeta = 0$, and $c_\infty = c_\mu \geq \nu_0$. We get therefore treat the case **(i)** as a special case of case **(ii)** to which we now confine our attention. We have the following cases.

- Case 1 $c_\infty = \nu_0$. In this case, following Lemma 4.2, (MPCC(c)) has a unique multiplier, and the result from [5] applies to give superlinear convergence of (x^q, ζ^q) to $(x^*, 0)$.
- Case 2 $c_\infty > \nu_0$ and $\zeta^q + \delta\zeta^q = 0$ for all $q \geq q_0$. Then the solution obtained of (QPMC(c)) is a also a solution of (QPX) which is feasible for all $x = x^q$. The result claimed follows from [18].
- Case 3 $c_\infty > \nu_0$ and $\zeta^{q_0} + \delta\zeta^{q_0} > 0$. Then using Lemma 4.4, part (i) it follows that the point $x^{q_0+1} = x^{q_0} + d^{q_0}$, where d^{q_0} are the components of the solution of the quadratic program (QPMC(c)) corresponding to x at (x^q, ζ^q), satisfies the complementarity constraints and that $x_{k1}^{q_0+1} = 0$, $k = 1, 2, \ldots, n_c$. Using now Lemma 4.4, part (ii), for all $q \geq q_0 + 1$, we obtain that x^q satisfies $x_{k1}^q = 0$, $k = 1, 2, \ldots, n_c$ for $q \geq q_0 + 1$ and that $\zeta^q = 0$ for $q \geq q_0 + 2$. The superlinear convergence result again follows from [18].

The proof is complete. $\diamond$

We have the following observations:

1. We do not prove here the more desirable result that if x^q is sufficiently close to x^* we obtain superlinear convergence to x^* or of (x^q, ζ^q) to $(x^*, 0)$. The difficulty is that, unless we have an estimate for ν_0, the subsequent iterates may find themselves far away from x^* once the elastic mode is entered. In effect, if the penalty parameter c_∞ is too small, the iterates may even be unbounded, even if the objective function is bounded on the feasible set. Such an adverse outcome can be prevented only by a global convergence result that will be the result of future research.
2. From Lemmas 4.2 and 4.3 we have that $||\zeta^q|| = O(||d^q|| + |d_\zeta^q|)$, so the c_∞ update rule from Table 4.1 will not be triggered, for x^q sufficiently close to x^* and for $c_\infty > \nu_0$. **Therefore, the update rule does not interfere**

with super linear convergence. On the other hand, it is also clear that, if $c_\infty < \nu_0$, then $\zeta^q > 0$ and the rule will eventually be triggered assuming that the iterates approach a stationary point. The test we use here is important because we do not spend an infinite amount of steps with an inappropriate c_∞. The complete study of an appropriate rule should also involve global convergence issues, since one possible outcome of the penalty parameter adaptation rule is to obtain unbounded iterates.

3. The appropriate value of c_μ is ν_0 (or a slightly larger value) which of course cannot be determined unless we specifically use the MPCC structure. A general NLP approach cannot be guaranteed to succeed in determining the appropriate value for c_∞ through c_μ by looking at the multipliers of (QPMC(c)) alone, since the Lagrange multiplier set of (MPCCS) is unbounded, whereas such rules are based on the assumption of boundedness of the Lagrange multiplier set at the solution x^* [2]. Whether c_μ can be adaptively defined is a matter for future research. For NLP implementations, a user-defined value for c_μ is considered to be an acceptable approach [24].
4. Note that the only meaningful difference between the proofs outlined here and the ones in [18], once the Lemma 4.4 has been established, is the one involve in the case where $c_\infty = \nu_0$, for which the well-known result from [5] has been invoked. Our contribution to this class of super linear convergent results has been essentially to show that, for the case $c_\infty > \nu_0$, the elastic mode approach, as presented in Table 4.1, will force the algorithm to choose points that are either always feasible or always satisfy the complementarity constraints and in so doing we do not have to apriori assume that this holds, as it was done in [18].

A similar result holds for the case where we use only Assumption **[A4]**, rather than its stronger version **[A4a]** that is used in [18]. The only weakening we have to do of the result is that we have to assume convergence of the sequences from the outset, rather than assuming only that the sequences find themselves in a sufficiently small neighborhood of the respective limit points. Following techniques from [45, 44], this assumption could probably be dropped at the cost of substantially complicating the analysis.

THEOREM 4.6. *Assume that Assumptions* **[A1]**–**[A4]***, hold near a solution x^* of (MPCCS). Assume that either*

(i) The elastic mode is never invoked, and the algorithm uses at q_0, for the purpose of constructing the matrix $\tilde{W}$, an estimate of the Lagrange multiplier that is sufficiently close to a multiplier of (MPCCS) and that $x^q \to x^*$.

(ii) The elastic mode is invoked at q_0 with $c_\infty \geq \nu_0$ and at all subsequent iterates, and the algorithm uses at q_0, for the purpose of constructing the matrix $\tilde{W}$, an estimate of the Lagrange multiplier that is sufficiently close to a multiplier of (MPCC(c)), and a ζ^{q_0} that is sufficiently close to 0, and that $(x^q, \zeta^q) \to (x^*, 0)$.

Then x^q converges to x^* superlinearly in case **(i)** *and (x^q, ζ^q) converges to $(x^*, 0)$ superlinearly in case* **(ii)**.

Proof Much as in the proof of Theorem 4.5, the proof of the case **(i)** can be reduced to the proof of the case **(ii)**. Also, the conclusion of the case **(ii)** where $c_\infty = \nu_0$ follows by the same argument as in Theorem 4.5. So we concentrate on proving **(ii)** for the case $c_\infty - \nu_0 = \gamma_0 > 0$.

We will complete the proof by considering the following cases.

Case 1. For all iterations, we have that $\zeta^q + d_\zeta^q = 0$, but the relation $x_{k1}^q = 0$, $k = 1, 2, \ldots, n_c$ does not hold at any iteration q.

Then on one hand, from Assumption [**A4**] and Lemmas 4.2 and 4.3 we must have that $\widehat{\mu}_{k1}^{q+1} > 0$ for $k = 1, 2, \ldots, n_d$. Assume that, on the other hand, we have that $\widehat{\mu}_{k1}^{q+1} > 0$ for some iteration index q and all k, $n_d + 1 \leq k \leq n_c$. Then, from the optimality conditions of (QPMC(c)) we must have that $x_{k1}^{q+1} = 0$, $k = 1, 2, \ldots, n_c$. This is a contradiction with the assumption at the beginning of this case.

So there must be a k such that $\widehat{\mu}_{k1}^{q+1} = 0$. Without loss of generality for the following argument, we assume that

$$\widehat{\mu}_{n_c,1}^{q+1} = 0 \tag{4.12}$$

Take now j, $n_{l_i} + 1 \leq j \leq n_i$. Since we have that $\zeta^q + d_\zeta^q = 0$, it follows that the double inequality constraint $-\zeta - d_\zeta \leq h_j + \nabla_x h_j(x) d \leq \zeta + d_\zeta$ of (QPMC(c)) will have both sides active, with corresponding Lagrange multipliers $\widehat{\pi}_j^{-q+1}$ and $\widehat{\pi}_j^{+q+1}$. An immediate consequence of Assumption [**A2**] is that $\widehat{\pi}_j^{-q+1} - \min\{\widehat{\pi}_j^{-q+1}, \widehat{\pi}_j^{+q+1}\}$ and $\widehat{\pi}_j^{+q+1} - \min\{\widehat{\pi}_j^{-q+1}, \widehat{\pi}_j^{+q+1}\}$ are also Lagrange multipliers of the same respective constraints. Note that the expression of $\widehat{W}$ at the next iteration $q+1$ does not change when using the new multipliers. We can therefore assume without loss of generality that

$$\widehat{\pi}_j^{-q+1} \widehat{\pi}_j^{+q+1} = 0. \tag{4.13}$$

Let now $\widehat{\Pi}^*$ be the Lagrange multiplier of (MPCC(c)) that is closest to

$$\widehat{\Pi}^{q+1} = \left(\widehat{\nu}^{q+1}, \widehat{\pi}^{q+1}, \widehat{\pi}^{-q+1}, \widehat{\pi}^{+q+1}, \widehat{\mu}_1^{+q+1}, \widehat{\mu}_2^{+q+1}, \widehat{\eta}^{q+1}, \widehat{\theta}^{q+1} \right).$$

Define now f_j^*, $j = n_d + 1, \ldots, n_c$, and a^* by subtracting the respective components of the fundamental multiplier from the components of $\widehat{\Pi}^*$, as in (4.5).

It then follows from Lemma (4.2) and Lemma (4.3) that

$$|\widehat{\mu}_{n_c,1}^{q+1} - \widehat{\mu}_{n_c,1}^*| \leq c_2 \left(||x^q - x^*|| + \zeta^q \right).$$

From equations (4.12) and (4.5) it then follows that

$$0 \leq -a^* x_{n_c 2}^* \leq \widehat{\mu}_{n_c,1}^* \leq c_2 \left(||x^q - x^*|| + \zeta^q \right),$$

which, in turn, implies that

$$0 \leq a^* \leq \frac{c_2}{-x_{n_c 2}^*} \left(||x^q - x^*|| + \zeta^q \right).$$

By a similar rationale, we get that (4.13) implies that, for any j such that $n_{l_e} + 1 \leq j \leq n_c$ we have that

$$f_j^* \leq c_2 \left(||x^q - x^*|| + \zeta^q \right).$$

Using now (4.5) we get that

$$\widehat{\Pi}^{q+1} \rightarrow (\Pi^*, \gamma_0),$$

where Π^* is the fundamental multiplier whose components are defined in (FMC).

Since the sequence of Lagrange multipliers is convergent, it follows from Lemma 4.2 by an argument similar to the one in [5], that is also discussed in [45, 44], that the sequence (x^q, ζ^q) is superlinearly convergent to (x^*, ζ^*). The proof of this case is complete.

Case 2 We have that $\zeta^q + d_\zeta^q = 0$, as well as $x_{k1}^q = 0$, for $1 \leq k \leq n_c$ at iteration $q = q^0$. Then, from Lemma 4.4 we have that $\zeta^q + d_\zeta^q = 0$, as well as $x_{k1}^q = 0$, for $1 \leq k \leq n_c$ for all iterations $q \geq q_0 + 1$. It then follows that the solution of (QMPC(c)) in the x component coincides with the solution of (QPR), since on the feasible set of (QPR) we have that $d^T W d = d^T \tilde{W} d$ with the notations of Lemma 4.4. Subsequently, the steps produced by the algorithm in the x component are identical to the steps produced by the algorithm applied to the nonlinear program

$$\begin{array}{rlcll} \text{minimize} & f(x) & & & \\ \text{subject to} & g_i(x) & \leq & 0 & i = 1, 2, \ldots, n_i \\ & h_j(x) & = & 0 & j = 1, 2, \ldots, n_j \\ & x_{k1} & = & 0 & k = 1, 2, \ldots, n_c \\ & x_{k2} & \leq & 0 & k = 1, 2, \ldots, n_c \end{array}$$

Since, from Assumptions **[A2]** and **[A3]** this program satisfies LICQ and SOSC at x^*, the fact that x^q converges superlinearly to x^* follows from [12, 38]. The result follows since $\zeta^q = 0$.

Case 3 We have that $\zeta^q + d_\zeta^q > 0$, at iteration $q = q_0$. It then follows from Lemma 4.4 that $x_{k,1}^{q+1} = 0$, for $1 \leq k \leq n_c$. Using Lemma 4.4 again, we obtain that $\zeta^q + d_\zeta^q = 0$, as well as $x_{k1}^q = 0$, for $1 \leq k \leq n_c$ for all iterations $q \geq q_0 + 2$. The conclusion follows from the preceding case. The proof is then complete. $\diamond$

Acknowledgments. Thanks to Stefan Scholtes for pointing out that the proof of Corollary 2.3 also stems from [40, Theorem 4]. Thanks to Alexander Shapiro for pointing out reference [5]. I am particularly grateful to my colleague Sven Leyffer for providing the AMPL examples and guidance for the numerical results and to Steve Wright for pointing out the result [45, Theorem A1]. Finally, I acknowledge the detailed comments from the referees and the editor, Danny Ralph, that have prompted the development of Section 4, and have markedly improved the quality of the results. This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract W-31-109-ENG-38. This work was also supported by award DMS-9973071 of the National Science Foundation.

REFERENCES

[1] M. Anitescu, *Degenerate nonlinear programming with a quadratic growth condition*, SIAM Journal of Optimization, 10(4) (2000), pp. 1116–1135.

[2] ———, *On the rate of convergence of sequential quadratic programming with nondifferentiable exact penalty function in the presence of constraint degeneracy*, Mathematical Programming, 92 (2002), pp. 359–386.

[3] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.

[4] ———, *Nonlinear Programming*, Athena Scientific, Belmont, Massachusetts, 1995.

[5] F. Bonnans, *Local analysis of Newton-type methods for variational inequalities and nonlinear programming*, Applied Mathematics and Optimization, 29 (1994), pp. 161–186.

[6] F. Bonnans and A. Ioffe, *Second-order sufficiency and quadratic growth for nonisolated minima*, Mathematics of Operations Research, 20(4) (1995), pp. 801–819.

[7] F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York, 2000.

[8] J. V. Burke, *A sequential quadratic programming method for potentially infeasible mathematical programs*, Journal of Mathematical Analysis and Applications, 139 (1989), pp. 319 – 351.

[9] ———, *A robust trust region method for constraint nonlinear programming problems*, SIAM Journal on Optimization, 2 (1992), pp. 325–347.

[10] J. V. Burke and S.-P. Han, *A robust sequential quadratic programming method*, Mathematical Programming, 43 (1989), pp. 277 – 303.

[11] R. J. Caron, H. J. Greenberg, and A. G. Holder, *Analytic centers and repelling inequalities*, European Journal of Operations Research, 143(2) (2002), pp. 268–290.

[12] A. Conn, N. Gould, and P. Toint, *Trust-region methods*, SIAM, Philadelphia, 2000.

[13] F. Facchinei, *a robust recursive quadratic programming algorithm with global and super linear convergence properties*, Journal of optimization theory and applications, 92 (1997), pp. 543–579.

[14] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York, 1983.

[15] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, Chichester, 1987.

[16] R. Fletcher and S. Leyffer, *Solving mathematical programs with complementarity constraints as nonlinear programs*, Optimization Methods And Software, (2004), p. To appear.

[17] ———, *Nonlinear programming without a penalty function*, Mathematical Programming, (to appear).

[18] R. Fletcher, S. Leyffer, S. Sholtes, and D. Ralph, *Local convergence of SQP methods for mathematical programs with equilibrium constraints*, Tech. Rep. NA 210, University of Dundee, Dundee, UK, 2002.

[19] R. Fletcher, S. Leyffer, and P. L. Toint, *On the global convergence of a filter-SQP type algorithm*, Numerical Analysis Report NA/197, Department of Mathematics, University of Dundee, UK, 1999.

[20] R. Freund, R. Roundy, and M. Todd, *Identifying the set of always-active constraints in a sytem of linear inequalities*, working paper, Sloan School of Management, MIT, Cambridge, MA, 1985.

[21] M. Fukushima, Z.-Q. Luo, and J.-S. Pang, *A globally convergent sequential quadratic programming algorithm for mathematical programs with linear complementarity constraints*, Computational Optimization and Applications, 10 (1998), pp. 5–34.

[22] M. Fukushima and P. Tseng, *An implementable active-set algorithm for computing a B-stationary point of a mathematical program with linear complementarity constraints*, SIAM J. of Optimization, (To appear).

[23] J. Gauvin, *A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming*, Mathematical Programming, 12 (1977), pp. 136–138.

[24] P. E. Gill, W. Murray, and M. A. Saunders, *User's guide for SNOPT 5.3: A fortran package for large-scale nonlinear programming*, Report NA 97-5, Department of Mathematics, University of California, San Diego, 1997.

[25] X. Hu and D. Ralph, *Convergence of a penalty method for mathematical programming with complementarity constraints*, Journal of Optimization Theory and Applications, (to appear).

[26] A. Ioffe, *Necessary and sufficient conditions for a local minimum.3: Second order conditions and augmented duality*, SIAM Journal of Control and Optimization, 17(2) (1979), pp. 266–288.

[27] ———, *On sensitivity analysis of nonlinear programs in Banach spaces: the approach via composite unconstrained optimization*, SIAM Journal of Optimization, 4(1) (1994), pp. 1–43.

[28] H. Jiang and D. Ralph, *Smooth SQP methods for mathematical programs with nonlinear complementarity constraints*, SIAM Journal of Optimization, 10 (2000), pp. 779 – 808.

[29] S. Lucidi, *Results on a continuously differentiable exact penalty function*, SIAM Journal on optimization, 2 (1992), pp. 558–574.

[30] Z.-Q. Luo, J.-S. Pang, and D. Ralph, *Mathematical Programs with Complementarity Constraints*, Cambridge University Press, Cambridge, UK, 1996.

[31] Z.-Q. Luo, J.-S. Pang, D. Ralph, and S.-Q. Wu, *Exact penalization and stationarity conditions of mathematical programs with complementarity constraints*, Mathematical Programming, 34(2) (1986), pp. 142–162.

[32] O. L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York, 1969.
[33] O. L. Mangasarian and S. Fromovitz, *The Fritz John necessary optimality conditions in the presence of equality constraints*, Journal of Mathematical Analysis and Applications, 17 (1967), pp. 34–47.
[34] B. Murtagh and M. A. Saunders, *Minos 5.0 user's guide*, Technical Report SOL 83-20, Stanford University, Stanford, 1983.
[35] *The NEOS guide*. Available online at *http://www.mcs.anl.gov/otc/Guide*.
[36] J. Outrata, M. Kocvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints: Theory, Applications and Numerical Results*, Kluwer Academic Publishers, 1998.
[37] E. Polak, *Optimization*, Springer, New York, 1997.
[38] S. M. Robinson, *Perturbed Kuhn-Tucker points and rates of convergence for a class of nonlinear programming algorithms*, Mathematical Programming, 7(1) (1974), pp. 1–16.
[39] ———, *Generalized equations and their solutions, part ii: Applications to nonlinear programming*, Mathematical Programming Study, 19 (1980), pp. 200–221.
[40] H. Scheel and S. Scholtes, *Mathematical programs with equilibrium constraints: Stationarity, optimality and sensitivity*, Mathematics of Operations Research, 25(1) (2000), pp. 1–25.
[41] S. Scholtes, *Convergence properties of a regularization scheme for mathematical programs with complementarity constraints*, SIAM Journal of Optimization, 11(4) (2001), pp. 918–936.
[42] A. Shapiro, *Sensitivity analysis of nonlinear programs and differentiability properties of metric projections*, SIAM Journal of Control and Optimization, 26(3) (1988), pp. 628–645.
[43] J. Telgen, *Minimal representation of convex polyhedral sets*, Journal of Optimization Theory and Applications, 38 (1982), pp. 1–24.
[44] S. J. Wright, *Superlinear convergence of a stabilized SQP method to a degenerate solution*, Computational Optimization and Applications, 11 (1998), pp. 253–275.
[45] ———, *Modifying SQP for degenerate problems*, SIAM Journal on Optimization, 13(2) (2002), pp. 470–497.

5. Response to the comments of the referees. .

5.1. Referee 1. The section 4 was vastly changed. In particular, its size has been reduced by more than 2 pages even though a result that uses weaker assumptions about the multipliers than [18] was added. This has been done by a change of notation (see assumption about signs of multipliers at the beginning of Section 4) and a change in the proof strategy, that revolves around Lemma 4.4. I believe the Section 4 is much more readable, which is what the referee requested.

5.2. Referee 2. .

1.1-1.2 The origin of the choices in question is now pointed out to [3, 4].

1.3 A model update is present in the algorithm defined in Section 4. The fact that it does not interfere with superlinear convergence is also discussed.

1.4 Note that the algorithm in Table 3.1 was added following the comment of the first referee about defining a specific algorithm about which the linear convergence claim is made.

1.5 Done.

1.6
- The respective phrase was corrected to read "... elastic mode, *such as* the one in SNOPT ...". Note that we regard the elastic mode as one way to relax the problem independent of what is used to solve the subproblem. In a paper that is the offshoot of the former global convergence part of the first version we use interior point algorithms to solve the relaxed problem.
- The point is well taken. The respective statement was deleted.
- Note that the respective statement refers to an aposteriori observation, rather than an apriori assumption.

2.1 The connection and no loss of generality statement of the 2 forms is described in the second paragraph of Section 4.

2.2 Note that the algorithm in Table 3.1 was simply one for which we proved linear convergence under weak second-order assumptions. They both still use elastic mode.

2.3 The statement of MPEC-SOSC was corected to not include MPEC-LICQ.

2.4 I thank for the referee for that observation, for I now realize that the point was not clearly made. I have added Lemma 4.1 to make the point that such a c_∞ can, in effect, be chosen.

2.5 I do not disagree with the referee, but note that the aim of our paper is quite different. We are trying to prove superlinear convergence without explicitly dropping the complementarity constraint, which is a modeling decision. Pursuant to the requests of the first referee at the first round of reviews, I have tried to define a black-box NLP elastic mode and show that it results in superlinear convergence when applied to MPCC. Dropping the complementarity constraint cannot be a part of the black box approach since it contains an inner modeling decision that is aware of the MPCC structure. I agree that keeping the constraint is what makes the whole analysis more difficult, as it can also be seen in [18].

3.1 I have done my best to unify it.

3.2 Note that a minimizer of the infeasibility includes no information about the objective function, which is included in that definition through the critical cone over which the second-order condition is defined.

3.3 Note that the conditions in those references are substantially weaker than Robinson's conditions which most references tend to consider the standard ones.

3.4 The point is well-taken. This is now discussed in the paragraph following the one where (MPCC) is defined.
3.5 The issue is discussed in the paragraph following the definition of (RNLP).
3.6 The convention about gradients is discussed in the new second paragraph of Subsection 1.2.
3.7 The issue is discussed following (2.14)–(2.15).
3.8 The distinction is important for our superlinear convergence result. For our analysis, it is essential not to relax the variables that enter the complementarity constraints.